DOCUMENT RESUME

ED 026 137

PS 001 530

By-Hall, Vernon; And Others
Project Head Start Research and Evaluation Center, Syracuse University Research Institute. Final Report, November 1, 1967.
Syracuse Univ., N.Y. Research Inst.
Spons Agency-Office of Economic Opportunity, Washington, D.C.
Report No-OEO-4010
Pub Date 1 Nov 67
Note-148p.
EDRS Price MF-$0.75 HC-$7.50
Descriptors-Age Differences, Auditory Perception, *Behavioral Science Research, Complexity Level, Concept Formation, Fear, Grade 2, Grammar, Intelligence Quotient, Kindergarten Children, Memory, Paired Associate Learning, Parent Influence, Preschool Children, *Preschool Evaluation, Reactive Behavior, *Research Projects, Sentence Structure, *Technical Reports, Test Interpretation
Identifiers-*Project Headstart

This document describes the research activities of the Syracuse University Evaluation and Research Center for the year September 1, 1966 through August 31, 1967. This final report is organized on the basis of six research projects, which have been abstracted under the following titles and numbers: (1) Experiments in Grammatical Processing in Children (PS 001 531); (2) Acquisition and Transfer Differences Between Kindergarteners and Second-graders on Aurally and Visually Presented Paired-associates Using an A-B, A-C Design (PS 001 532); (3) Concept Identification Strategies (PS 001 533); (4) Fear and Attachment in Young Children (PS 001 534); (5) Evaluating Behavioral Change During a Six-Week Pre-kindergarten Intervention Experience (PS 001 535); and (6) Variables Affecting the Performance of Young Children on a Letter Descrimination Task (PS 001 009 or ED 020 797). (WD)

ED026137

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Project Head Start Research and Evaluation Center

Contract # OEO 4010
Final Report
November 1, 1967

William J. Meyer
Center Director

Cooperating Agency
Syracuse University
Research Institute
201 Marshall Street
Syracuse, New York 13210

Institutional Grants Officer
L. William Patchen
Assistant Director
Research Institute
Syracuse University

PS001530

TABLE OF CONTENTS

I. Introduction Page 1

II. Research Projects

1. Experiments in Grammatical Processing in Children. Page 2
   Murray S. Miron, Vernon Hall, Michael Mery

2. Acquisition and Transfer Differences Between Kindergarten and Second-graders on Aurally and Visually Presented Paired-associates Using an A-B, A-C Design. Page 22
   Vernon C. Hall

3. Concept Identification Strategies. Page 30
   William J. Meyer and David Hultsch

4. Fear and Attachment in Young Children. Page 44
   J. Conrad Schwarz

5. Evaluating Behavioral Change During a Six-week Pre-kindergarten Intervention Experience. Page 58
   Lois Hayweiser, David Massari, and W.J. Meyer

6. Variables Affecting the Performance of Young Children on a Letter Descrimination Task. Page 92
   Vernon C. Hall, Edward Caldwell, and Gwen Simpson

III. Appendix A

1. Adaptive Behaviors Page 99

2. STIM Guide to Interviewers Page 102

3. Math, Pre and Post-Test Page 105

I. Introduction

This report describes the research activities of the Syracuse University Evaluation and Research Center for the year September 1, 1966 - August 31, 1967. The report is organized in terms of the work of each of four investigators, Vernon Hall, William J. Meyer, Murray Miron and J. Conrad Schwarz.

As the reader proceeds through the content of the report, it will become clear that, with one exception, the children employed in the various studies are not from especially culturally disadvantaged environments. For the most part our available pool of research subjects are from a lower-middle class environment and are attending public schools in a suburban area adjacent to the city of Syracuse. It was necessary to use these children in our research because of the unavailability of Head Start programs in the community; the lack of adequate research space within the Pre-kindergarten Centers (supported by New York State with Title I funds from the office of Education); and, more seriously, a general feeling of misgiving among the teachers in the community about the usefulness of research programs.

Substantial progress was made this past year in remedying the subject situation. A series of meetings were held with the director of the Pre-kindergarten Program in which our mutual concerns were discussed. These discussions have resulted in a working relationship as evidenced by the recent approval of two proposed projects. Space remains a problem which will probably be resolved through the purchase of a mobile laboratory.

Ultimately, the Evaluation and Research Center would like to administer a research and demonstration center comprised of four to six classes. The distinct advantage to such a center is the freedom it provides for intervention research.

Control over subject selection and curriculum, for example, is readily possible when administrative responsibility rests with a research team.

II. Research Projects

1. Experiments in Grammatical Processing in Children

Murray S. Miron, Vernon Hall, Michael Mery

The experiments of this report are predicated on the assumption of a particular model of what it is that constitutes language and how it is that language is realized and understood as speech. Traditional models of the speaker/hearer have presumed that the user of language acts as a passive-filtering mechanism which serves to sort output or input units into molecular categories in canonical or Markovian fashion. The classic influences of the taxonomic grammars current up to the last decade, the deceptively obvious sequential ordering of speech in linear time increments, and the currency of psychological experiments on word frequency, stochastic influences in language and molecular reinforcement all conspired to encourage theories of language processing which did not distinguish between speech and language. Speech clearly is time ordered and uniformly is displayed as a series of concatenated units each occupying a linear increment of real time.[1] By analogy to the written forms of speech, these units were tentatively identified as word or sometimes even letter segments. The speaker of a language, and afterall nearly every hearer is a speaker, reflected upon the salience of word units and concluded therefrom that words had direct physical representation in speech.

[1] We may ignore, as the traditional approaches ignored, the simultaneous prosodic features (intonation, stress, etc.) which are imposed on top of these linear segments, if we recognize that such oversight is symptomatic of the taxonomic approaches which are under attack.

Most hearers and even some hearers who theorize about language have argued that words in speech are separated by energy dimunition points or pauses. Whatever the status of word units in language as a system may be, it is clear that no physical characterization of words is to be found in the acoustics of speech. It is a direct consequence of the traditional lack of separation in psychology of the systems of speech and language which engenders such erroneous observations.

By contrast there has recently emerged an explicit theory of language which sharply separates competence from performance, speech from language. This theory most closely identified with the pioneering monograph of Noam Chomsky (1957) attempts to explicitly characterize language as an hierarchically organized system of rules which permute syntactically molecular units into regularized structures of increasing abstraction and decreasing simularity to the terminal strings which are manifest in speech either as input or output of the system. The grammar of a language, in this view, is the deep lying set of regularities which account for, i.e., describe and explain, manifest structure. Understanding of any given utterance as a realization of the application of the linguistic system of rules, that is, as a well-formed utterance in a particular language; thus is viewed as the process of recreating the structural history of that utterance. Such a view presupposes an active model constructing, parsing approach to the decoding of language. Manifestly concatenated elements in speech must be tested against a permuted and abstract representation generated by the hearer seeking a derivational account of what is experienced in order to achieve understanding and thus assign meaning to that experience. Manifest content is thus distorted into a series of abstract concurrences of derived and perceived experiences. Further, the derived structure insofar as it predicts manifest content which may not as yet have been experienced, directs and controls perception in seeking confirmation of the hypothesized structural derivation. Speech, according to this model, is thus segmented into those units which correspond to predicted units by the assigned derivation. Whereas the concatenation model would predict a filtering of input experience into empty category slots, probabilisticly ordered in terms of their input access, this model superimposes structure upon input.

Several recent lines of investigation lend considerable credence to this active model of speech perception. Ladefoged and Broadbent (1960) in presuming such a model, reasoned that perceptual units imposed upon input experience should show a tendency to resist extraneous interruptions. From the viewpoint of guided experience, distortions of the manifest experience which are not relevant, i.e., are not predicted by the derived modes which attempts to organize that experience, should either be ignored or at least resisted. Such observations have been well established for other forms of human experience, visual illusions providing the most dramatic of these. With regard to speech, the predicted resistance to irrelevant distortion could, as these investigators argued, serve to identify the superimposed segmental boundries being employed by the listener. To this end, Ladefoged and Broadbent added an extraneous, non-speech click to each of a series of sentences and digit strings. The S̲s were instructed to identify the word which provided the locus for the extraneous sound. If this extraneous signal fell within a perceptual segment, its accurate location should be disturbed according to the size of the perceptual units which were being employed in the organization of the speech input. The investigation confirmed that errors in click location were indeed larger for sentential than non-sentential material. This finding at least grossly suggested that sentences were not being processed as discrete word units as is necessarily the case with random strings of digits. The presence of viable phrase units was at least plausible.

Any more refinement in these results requires a theoretical model which explicitly predicts the phrase segmentations of the speech stream on the basis of some inherent grammatical structure. One such model is that of Immediate Constituent Analysis elaborated some time ago by Rulon Wells (1947). In this model the surface structure of a given sentence is examined for constituent elements of decreasing complexity which by reference to substitutability and expansion in other utterances can be said to constitute a meaningful collocation of units. Thus, given the utterance: The king of England opened parliament, IC analysis would decompose the sentence first into the sub-units of (The king of England) and (opened parliament) on the grounds that the first constituent can be viewed as an expansion of a substitutable element such as (John) and the second constituent as an expansion of the substitutable element (worked) as in the complete utterance (John worked). The two constitutes thus isolated can be further analyzed into lower order constituents of (The king), (of England) and (opened) and (parliament). The remaining constitutes when analyzed into (The), (king) (of) and (England) finally reduce the sentential constitute into its elemental word constituents. It should be clear from the Polish notation that the final constituent structure:

(((The) (king))((of)(England)))((opened)(parliament)))

is hierarchally grouped into decreasingly complex constitutes which terminate in word elements of the total sentence. Further analysis is possible, e.g., the constituent (opened) might have been identified as the constitute composed of (open) and (-ed). For the present purposes such morphemic subdivision will not be employed. It should also be clear that each of the constitutes can be identified

PS001530

with one or another of the grammatical class markers which are the basis for their identification. Thus the first constitute (The king of England) is separable precisely because it is a substitutable expansion of the grammatical marker: subject. (Of England) in turn is a prepositional phrase which is functionally equivalent to all expansions of such phrases. Further, the IC analysis informs us that (of England) belongs to the constitute (The king of England) and consequently is a prepositional phrase used in adjectival function, precisely as (England's) (king) might have been used.

It is not, however, clear that the IC analysis can be a precise description of the underlying perceptual segmentations the hearer is presumed to use. IC analysis has the potential inadequacy of dealing with manifest structure as if such structure revealed underlying regularity. Its procedures are nonetheless reasonably explicit and easy of application.

Several experiments have in fact employed the IC model as a test of the segmentations presumed to occur in perception. Garrett (1964) and Fodor and Bever (1965) demonstrated that subject displacement of the perceived locus of a transient non-speech event (a click produced by condensor discharge superimposed upon the speech material) could reliably be predicted from a knowledge of the IC constitutes of sentential material. In general, the deeper the level of constituent boundary, the greater was the tendency for perceived location of the position of the click. The depth of an IC boundary is equivalent to the syntactic complexity of the phrase segment involved. Thus, to return to our earlier example, the constitute boundary between England and opened is the deepest level of IC segmentation exactly because it represents the largest and hence most complex division of the utterance which is possible of subdivision. On either side of this deepest boundary, additional minor level segmentations can be identified. In general, the Polish notation reflects depth of boundary as a function of the number of parentheses which reverse the left-to-right flow of organization. The structure ((A)(B))(C) differs from ((A)((B)(C)) in that the former has as its major constituent boundary the division between B and C and the latter between A and B. Both constructions differ from the construction ((A)(B))((C)(D)) in that the boundary between B and C of this construction is deeper than those of the former two. It should be noted that no more than an ordinal scale of "depth of structure" is implied by this model, no motivation is being offered for an interval measure of the construct. In fact, only Yngve (1961) has serviously suggested comparing different sentences with regard to boundary depth in connection with speech processes.

In discovering that click distortions were drawn to major constituent boundaries, the previous investigations had subtly confirmed the importance of linguistic segmentation in perception. The objection that major constituent boundaries may also be marked by events which might have drawn the click perception, events such as pause or intonation, was subsequently eliminated in a further study by Garrett, Bever and Fodor (1966).

Thus it must be presumed that the hearer organizes the input speech stream into linguistic segments and that extraneous events which cut these segments are perceptually deferred until processing reaches a completed unit of segmentation.

By way of summary, the following conclusions appear to be well established. (1) Brief speech distortions in the form of readily perceived clicks migrate from their objective locii when imposed upon sentential material to the nearest major syntactic boundary as defined by immediate constituent analysis. (2) Accurate location of the objective locii of click distortions in sentential material is significantly enhanced when the locus of the distortion is adjusted to correspond with the major syntactic boundary. (3) Displacement of perceived locii of distortions is significantly reduced when non-sentential material is employed. (4) Neither pause events nor intonation cues, although in themselves capable of displacing perceived locii of distortions, when controlled, are found to significantly alter the attractive power of the linguistic segmentation. (5) These results are consonant with an interpretation which assumes that active syntactic segmentation is a necessary part of speech perception and that displacement of distorting events in speech is a result of an attempt on the part of the hearer to preserve the unitary integrity of those segments.

All of the foregoing experimental results are based upon an adult population of subjects. There is sound a priori reason to believe that children might differ from adults with regard to these effects. In reductio ad absurdum, the prelinguistic child cannot be aware of segmentation in the speech stream and hence cannot be expected to attempt to preserve the integrity of such segments. Presumptively, the child proceeds by orderly stages from this pre-linguistic state through increasing linguistic competence involving progressively longer segmentation units of his input (cf Bellugi, 1967). Complex syntactic constructions beyond his competence stage should be either incorrectly segmented or not segmented at all. The click displacement procedure thus offers the investigator a subtle means of trying to trace the longitudinal development of linguistic competence in the child.

Several complications arise, however. First, for very young children the testing conditions themselves are complicated enough to make instructioning the child a potential obstacle. Second, the child appears to be much more sensitive to lexical difficulty than is the adult. In theory, given a particular syntactic construction with its appropriate IC segmentation, vocabulary (lexical) changes in no way modifies that structure. It is for this reason that the Carroll jabberwocky, "The slithy toths did gyre and gimble." is comprehensible. Consider the following three well formed English utterances.

1. (((The) (friends))((helped)((by)(Jane))))((tried)(hard))((for)((the)(prize)))
2. (((The)(reporters))((assigned)((to)(George))))((drove)(furiously))((to)((the) (station)))
3. (((The)(analysists))((allocated)((to)(Omar))))((schemed)(feverishly))((for) ((the)(renumeration))))

As is apparent from the segmentations, all three utterances are instances of an identical construction. Unfortunately, no experimentation had manipulated this variable for adults or children.

Ancillary Activity

It seemed clear to the investigators that several pertinent problems had to be resolved before an actual experiment could be devised. First, despite the fact that the senior author had been involved in the click experimentation from the time of Garrett's original experiments (Garrett, 1965) no firm procedural rules had been worked out for such crucial matters as mode of presentation of the distorting signal, type of distortion, and amplitude of distortion. Although the earlier investigators tacitly assumed that such variables were not crucial to the displacement effect, no empirical evidence for such an assumption had been gathered.

Consequently, the present investigators began with a series of pilot studies which manipulated the mode of presentation of the transient distortion. In all, five procedural conditions were employed. In all instances, the superimposed distortion was a 60 msec., 1K Hz. tone burst adjusted to the RMS voltage level of the peak vowel of the utterance sample. Oscillographic tracings of the tone burst indicated that the rise and decay time was sufficiently fast to produce a square wave trace of reasonable sharpness. Listening judgments confirmed that the tone burst was readily apprehended, was free of switching distortion and was of sufficient duration to give rise to the perception of a distinctive periodic character differing markedly from the speech events upon which it was superimposed.

The five procedural conditions, in their order of presentation to Ss, were as follows.

1. Speech in one ear, click only to the other ear.
2. Speech in one ear, click plus white noise in the other ear.
3. Speech in both ears, click multiplexed with speech in the other ear.
4. Speech in both ears, click multiplexed with speech in both ears.
5. Speech in both ears, no click.

Ten adult Ss, tested two at a time, were exposed to all five conditions. They were given printed versions of each of ten recorded sentences and instructed to expose the printed version immediately after listening to its spoken rendition, indicating thereon with a slash mark the precise locus of the extraneous signal. Ss were told that the tone-click could occur anywhere in the sentence, either on a particular word or between words. In addition, they were falsely informed that the repetitions of the sentences for the various listening conditions might differ and that they were to indicate whether or not the sentence had been modified by a same/different judgment. The latter procedure was adopted in order to assure that the multiple repetitions of the same sentence would receive equally careful attention.

The stimulus materials were devised such as to recreate the structural form of those employed by Garrett, Bever and Fodor (1966) while manipulating lexical difficulty at two levels. Table 1 displays these sentences with their numbering

TABLE 1 - Adult Pilot Stimulus Materials Paralleling Those Employed by Garrett, _et al._, (1966) with Manipulated Lixical Difficulty

1a(1). IN ORDER TO CATCH HIS TRAIN, JO*HN DROVE QUICKLY TO THE STATION.

1a(2). IN ORDER TO ELUCIDATE HIS POSTULATE, LOBES*CHEVSKY TOILED DILIGENTLY IN HIS ARBORETUM.

1b(1). THE FRIENDS HELPED BY JANE TRIED HA*RD FOR THE PRIZE.

1b(2). THE ANALYSTS ALLOCATED TO OMAR SC*HEMED FEVERISHLY FOR THE RENUMERATION.

2a (1). IN HER WISH FOR CA*NDY MARY WAS VERY SAD.

2a(2). IN HER EXPECTATION OF NUPT*IALS, TRIBLENKA WAS SURELY SOPHISTICAL.

5a(1). AS A GOOD PART OF HIS NEW FRIEND'S PR*ESENT, THE BOY WAS GIVEN A QUARTER.

5a(2). AS A CONCOMITANT CONSEQUENCE OF THEIR RECENT INNOVATION'S IMP*ACT, THE CARTEL WAS ALLOTTED A DISPENSATION.

5b(1). THE OLD SHERIFF WHOSE IDEAS STILL GREATLY AFFECT THE WEST WAS*GIVEN A PRIZE.

5b(2). THE OCTOGENERIAN ADJUDICATOR WHOSE BRIEFS STILL GREATLY SWAY THE B*AR WAS AFFORDED AN ACCOLADE.

*Indicates objective locus of tone-burst.

keyed to the comparable sentence form used by Garrett, *et al*.

The distribution of click displacements confirmed that found by Garrett, *et al*. Lexical difficulty, as predicted, did not exert an influence on the degree or occurrence of displacements. The first of the procedural conditions matching that employed in the earlier investigations (speech in one ear, tone-burst in the other) produced the greatest magnitude of perceived click displacement. The second condition which added white noise to the click ear-channel produced the second greatest total amount of displacement. Despite the fact that this second condition did not produce as much displacement as the first, it was this noist plus click condition which was adopted for all subsequent testings. The quiet dichotic listening channel of the first condition introduced the possibility of an extraneous attention factor in the task. Listening under such a condition subjectively conditions a hearer to expect sound only in one ear. The short tone-burst which punctuates an otherwise quiet channel may very well catch the subject unready for sound in the unused ear. Displacement of the click locus under such a condition could be argued to be contaminated by readiness factors which are extraneous to the segmentation effects. The continuous white noise condition with its brief overlaid tone burst holds the distortion channel open. In addition, the quiet channel of condition one is necessarily a relative determination. Noise is, of course, present, it is merely undefined in character and at a low level. Flaws in the recording tape, random cathode emissions from the tube components of the amplifier and stray external electromagnetic radiation all produce noise peaks which conceivably could be mistaken for the experimentally manipulated click. Sound experimental practice thus appeared to dictate the use of the determinant noise condition over that employed by the other investigators. Parenthetically, it should be noted that it was for similar reasons that the 1K Hz. tone rather than the condensor produced discharge was adopted in these experiments. The periodic character of the 1K tone makes it a distinct event which can readily be distinguished from extraneous or inherent noise.

Thus encouraged that the experimental effects could be replicated, that the acoustical conditions were appropriate and that lexical difficulty was not a controlling variable for adults, additional pilot data was collected on a small sample of child Ss.

On the grounds that the Illinois test of Psycholinguistic Abilities (McCarthy and Kirk, 1961) provided a source of stimulus materials which had both established norms and high face validity for a task to be used with children, the test was searched for appropriate sentential material. Two subtests of the ITPA were selected, the Auditory-Vocal Association Test and the Auditory Decoding Test. The ITPA arranges subtest items in order of difficulty as determined from the standardization samples employed. Within a single subtest, however, the syntactic form of the items remains constant. Level of difficulty is thus controlled on the ITPA by changes in the lexical difficulty level, the variable still needing assessment for the child subject samples. The items of the Auditory-Vocal Association subtest are of the form:

((I))((sit)((on)((a)(chair))))((I))((sleep)((on)((a)(____))))

and those of the Auditory Decoding subtest:

((Do))((cars)(cry))

As can be seen, the form of the items provides for a single, quite deep syntactic boundary in each of the sentences. Both subtests require that the subject make a single lexical response which when compared to the norms for that item can be scored as either correct or incorrect. Thus it is possible to confirm understanding of the utterance and simultaneously assess the change in norm levels with the introduction of the additional task of click identification.

The non-sentential control condition to be used for comparison purposes was drawn from the items of the Auditory-Vocal Sequencing subtest which employs haphazardly ordered digits strings.

In order to assess the best response condition for children it was decided that two forms of judgment would be manipulated. In the first condition, Ss were instructed to identify the position of the tone-burst directly. In the other condition, Ss were given the task of indicating whether or not the tone-burst occurred at the same or a different position in the repeated sentence. Locii identification was used for the Auditory-Vocal Association items and same/different judgments for the Auditory-Vocal Sequencing and Auditory Decoding items. Table 2 displays the full complement of stimulus materials employed in these initial testings. All Ss were exposed to the full set of materials in the order: Auditory Vocal Association, Auditory Decoding and Auditory-Vocal Sequencing. Order of item presentation within a subtest followed that of normative difficulty level. It will be observed that for the locii identification task, half of the items contain no tone-burst at all, systematically counter-balanced against the remaining half for which there was a tone-burst. This procedure enabled a separate scoring of simple click detection in case the child could recognize the presence of a click but could not (or would not) venture an identification of its position.

The same/different judgments, similarly counter-balanced the occurrences of identical or changed position of the click instances.

TABLE 2 - Child Pilot Stimulus Materials
Auditory-Vocal Association Test (ITPA, p. 41, nos. 1-20)

1. I sit on a chair; I sleep *on a ____.
2. I eat from a plate; I drink from a ____.
3. A bird flies in the air; a fish swims in the ____.
4. I hit with my hands; I kick with my ____.
5. John is a boy; Mary *is a ____.
6. A scissors cuts; a pencil ____.
7. I cut with a saw; I pound *with a ____.
8. Soup is hot; ice cream is ____.
9. A red light says stop; a green light says ____.
10. During the day we're *awake; at night we ____.
11. I eat with a spoon; I *cut with a ____.
12. On my hands I have fingers; on my feet I have ____.
13. A boy runs; an old *man ____.
14. Cotton is soft; stones are ____.
15. An explosion is loud; a whisper is ____.
16. Mountains are *high; valleys are ____.
17. A man may be a king; a woman may *be a ____.
18. A pickle is fat; a pencil is ____.
19. Coffee is bitter; sugar *is ____.
20. Iron is heavy; feathers are ____.

---

*Objective placement of tone-burst

(TABLE 2 - cont'd.)

Auditory Decoding Test (ITPA, p. 58)

	STIMULUS STRING	CONDITION
1.	D$\overset{*}{\text{o}}$ you smoke?	same
2.	D$\overset{*}{\text{o}}$ you r$\overset{\#}{\text{u}}$n?	different
3.	D$\overset{\#}{\text{o}}$ you f$\overset{*}{\text{l}}$y?	different
4.	Do you b$\overset{*}{\text{a}}$rk?	same
5.	D$\overset{*}{\text{o}}$ babies eat?	same
6.	D$\overset{*}{\text{o}}$ bicycles dr$\overset{\#}{\text{i}}$nk?	different
7.	Do ap$\overset{*}{\text{p}}$les fly?	same
8.	D$\overset{\#}{\text{o}}$ dresses dr$\overset{*}{\text{i}}$ve?	different
9.	D$\overset{*}{\text{o}}$ bananas tel$\overset{\#}{\text{e}}$phone?	different
10.	Do balls bo$\overset{*}{\text{u}}$nce?	same
11.	D$\overset{*}{\text{o}}$ eagles paint?	same
12.	D$\overset{\#}{\text{o}}$ goats e$\overset{*}{\text{a}}$t?	different
13.	Do pinc$\overset{*}{\text{u}}$shions cheer?	same
14.	D$\overset{*}{\text{o}}$ children c$\overset{\#}{\text{l}}$imb?	different
15.	D$\overset{\#}{\text{o}}$ lanterns sh$\overset{*}{\text{i}}$ne?	different
16.	Do daughters ma$\overset{*}{\text{r}}$ry?	same
17.	D$\overset{*}{\text{o}}$ dials yawn?	same
18.	D$\overset{*}{\text{o}}$ barometers congra$\overset{\#}{\text{t}}$ulate?	different

---

*First position of tone-burst in change conditions or single locus in same condition.

#Second position of tone-burst in change conditions.

(TABLE 2 - cont'd.)

Auditory-Vocal Sequencing Test (ITPA, p. 50)

STIMULUS STRING	CONDITION
1. $\overset{*}{1}\ 2$	same
2. $\overset{*}{9}\ \overset{\#}{6}$	different
3. $\overset{\#}{5}\ \overset{*}{2}\ 1$	different
4. $6\ \overset{*}{8}\ 9$	same
5. $9\ 7\ \overset{*}{6}$	same
6. $\overset{*}{4}\ 3\ \overset{\#}{4}$	different
7. $6\ \overset{*}{6}\ 1\ 6$	same
8. $6\ \overset{\#}{3}\ 5\ 8$	different
9. $\overset{*}{4}\ 4\ \overset{\#}{2}\ 4$	different
10. $5\ 7\ \overset{*}{4}\ 5$	same
11. $2\ 5\ 4\ 9\ \overset{*}{9}$	same
12. $6\ \overset{\#}{1}\ 6\ \overset{*}{3}\ 7$	different
13. $\overset{*}{4}\ 5\ 9\ 1\ 4$	same
14. $9\ \overset{*}{1}\ 7\ 5\ \overset{\#}{3}\ 3$	different
15. $\overset{\#}{8}\ 9\ 6\ 3\ \overset{*}{4}\ 8$	different
16. $6\ 3\ \overset{*}{9}\ 7\ 3\ 5$	same
17. $6\ 9\ 2\ 8\ \overset{*}{7}\ 9$	same
18. $3\ \overset{*}{8}\ 9\ 1\ 7\ \overset{\#}{5}\ 5$	different
19. $6\ 4\ 5\ \overset{*}{5}\ 8\ 4\ 1$	same
20. $8\ 2\ \overset{\#}{9}\ 4\ 7\ 5\ \overset{*}{3}$	different

---

*First position of tone-burst in change conditions or single locus in same condition.

#Second position of tone-burst in change conditions.

Unfortunately, given the intricate stimulus controls and the desirability of using the ITPA test items, no useful results were forthcoming. It was discovered that the kindergarten Ss could not sufficiently comprehend what it was that was required of them to complete the task. The Ss, their full attention focused upon correctly answering the sentence items; i.e., in comprehending the stimulus materials, were confused as to what was meant by a "click". That the difficulty was produced by the nature of the stimulus materials and not poor instruction is confirmed by their grasp of the task when the number sequences were employed. It was clear that the level of difficulty of these materials had pushed the children to the maximum information loading they could tolerate even for those items having lowest age norms. (Scoring norms for the ITPA go down to a language-age score of two years four months). Accordingly further testing with this stimulus sample was abandoned. The success obtained with the digit task, however, led the experimenters to believe that it was possible for the young Ss to identify the locus of the click in the speech stream, an important consideration if magnitude of displacement was to be measured. Thus, in the subsequent testings only the procedure involving identification of click locus was employed.

The experimenters are aware that up to this point in the research nothing of a substantive nature had been accomplished. Procedural issues, however, had been defined and solved. Other investigators had, in fact, attempted to use children as respondents in a task of this type and had failed to even progress to the meager point at which this ancillary research had left us.

Experimental Task

Stimulus Materials and Subjects. A total of eight practice utterances and 20 test items were devised such as to be well within the linguistic competence of each of two subject samples of eight children drawn from kindergarten and second grade classes of a local elementary school. The test materials are displayed in Figures 1 and 2. The eight practice items which always preceeded the test items were repeated twice, once at low reading speed and then again at normal rate. These eight utterances are displayed in Table 3.

It will be observed that each of the 10 test sentences has a corresponding non-sentence produced by randomly ordering the lexical items. Objective placement of the tone-bursts was on the identical lexical item in either version. This position is noted by the "0" line of the plots of Figures 1 and 2. As before, the tone-burst was a 60 msec., 1K Hz. signal. Administration of the test items was in a counter-balanced order which controlled for sentence or non-sentence position.

Procedure. Each subject was brought to the testing room by the experimenter and was seated at a small table opposite the experimenter. The subject was told that in a minute, he would put on the earphones and he would hear a voice saying some words. He would also hear a "click" or "beep" and was to raise his hand as soon as he heard it. This was done to insure that the child knew what the click or beep sound was. Both the subject and the experimenter then put on earphones and word group #1 was played. In case the child did not raise his hand when the click occurred, the word group was repeated until he did so.

The S was then told "Now, I'm going to play the same words again. This time, instead of raising your hand, tell me what word the man is saying when you hear the click. Tell me when the man is through speaking." Word Group #1 was then replayed. In case the S did not respond or replied with several words rather than one word, he was told "Was the man saying 'house', 'lion', 'mountain', 'chicken', or 'letter', (the words in the first item of Word Group #1). Word

TABLE 3 - Practice Stimulus Materials

1. House, lion, *mountain, #chicken, letter.
2. Table, *chair, desk, chimney, man.
3. One, #four, zero, nine, *two, three, seven.
4. Ten, *seventeen, twentyseven, #fiftysix, twentyeight, fiftynine.
5. The *house was on #top of the mountain.
6. The #chicken was *eating the grain.
7. The *letter was #written to the man.
8. The #teacher *thought it was Sunday.

---

*Location of tone-burst for slow reading.

#Location of tone-burst for normal reading.

Group #1 was repeated until the child responded with only one word. Most Ss did this one in one trial and all Ss did so within three trials. The next seven items were then played, all at the slower than normal speaking rate and the instructions were repeated in case any questions arose.

After the first eight word groups (slow rate) were played, the next eight were played. This section is the same as the first except it is spoken at a normal speaking rate. The subject was told "Now the man is going to tell you to do something. Do what he tells you to do and also tell me what the man is saying when you hear the click." These instructions were repeated before each of the twenty following items, ten items being commands requiring a motor response, When a nonsense string, of which there were also ten, was played, the child was first told "Here are some words. Just tell me what the man is saying when you hear the click." The full twenty items were then played and when completed, the subject's response to a particular item was recorded by the experimenter immediately after the response. Upon completion of the twenty items, the subject was returned to his classroom.

Results. The displacement data was analyzed by means of two separate analysis of variance models each employing three classification variables, viz., (1) grade at two levels, kindergarten and second, (2) word order at two levels, sentence and non-sentence orders and (3) utterance composition at 10 levels, i.e., the various lexical compositions. Both models of the analysis of variance were of the mixed type in which the three classification variables were assumed to be fixed, with subjects serving as random replications. The analyses are of the partially hierarchial type in that subject effects are necessarily nested within grade level but form complete replications for all levels of all other variables. "Never pool" procedures were practiced in the choice of error terms in all analyses.

The major two analyses were based upon separate scorings of the displacement data which either coded (1) direction of displacement as a signed value (left displacement from the objective locus scored as minus, right displacement as plus) or (2) magnitude of displacement as an absolute measurement. The former scoring procedure will henceforth be referred to as the algebraic scoring and the latter as absolute scoring. Consider the sample observation:

Stand up/and then $\overset{*}{\text{sit}}$ down.

where the slash mark, /, indicates the perceived location of the click placed upon the word *sit*. In scoring this observation by the algebraic criterion, the *S* would receive a score of -2.5, that is to say, one-half unit of displacement for each word and word-division unit from the objective position of the click. In absolute scoring, the *S* would receive a score of simply 2.5.

Tables 4 and 5 detail the variance contributions for the classification effects with respect to algebraic and absolute scores, respectively.

It will be observed that the two analysis confirm each other with regard to the significance of all effects with only the exception of the sentence by utterance interaction, (B)(C), it being significant when algebraic scores are considered and non-significant for absolute scores. Discussion of this difference will be deferred for the moment.

Sentence and utterance composition are clearly the two significant contributions controlling the observed variation in responding. The replicated effects under the two forms of scoring are most easily seen by reference to Figures 3 and 4. The plots for the combined groups scored algebraically or as absolute displacements are nearly identical. What discrepancy exists between the two plots is

TABLE 4 - Summary of Analysis of Variance of Algebraic Displacement Scores

Variance Source	SS	df	MS	F
Between Ss	240.49	15		
Grade (A)	12.02	1	12.02	-----
Ss within A (SwA)	228.47	14	16.32	
Within Ss	1515.04	304		
Sentence/Non-Sentence (B)	200.03	1	200.03	22.94*
(A) (B)	.90	1	.90	-----
(B) (SwA)	122.07	14	8.72	
Utterance Composition (C)	165.35	9	18.37	5.48*
(A) (C)	40.03	9	4.45	1.33
(C) (SwA)	422.62	126	3.35	
(B) (C)	114.27	9	12.70	3.86*
(A) (B) (C)	35.27	9	3.92	1.19
(B) (C) (SwA)	414.50	126	3.29	

---

* p less than that for alpha level .05

TABLE 5 - Summary of Analysis of Variance of Absolute Displacement Scores

Variance Source	SS	df	MS	F
Between Ss	235.52	15		
Grade (A)	5.52	1	5.52	-----
Ss within A (SwA)	230.00	14	16.43	
Within Ss	712.71	309		
Sentence/Non-sentence (B)	18.53	1	18.53	5.36*
(A) (B)	1.95	1	1.95	-----
(B) (SwA)	48.39	14	3.46	-----
Utterance/Composition (C)	54.52	9	6.06	2.49*
(A) (C)	29.43	9	3.27	1.35
(C) (SwA)	306.72	126	2.43	-----
(B) (C)	22.11	9	2.46	1.51
(A) (B) (C)	25.26	9	2.81	1.72
(B) (C) (SwA)	205.80	126	1.63	

---

* p less than that for alpha level .05

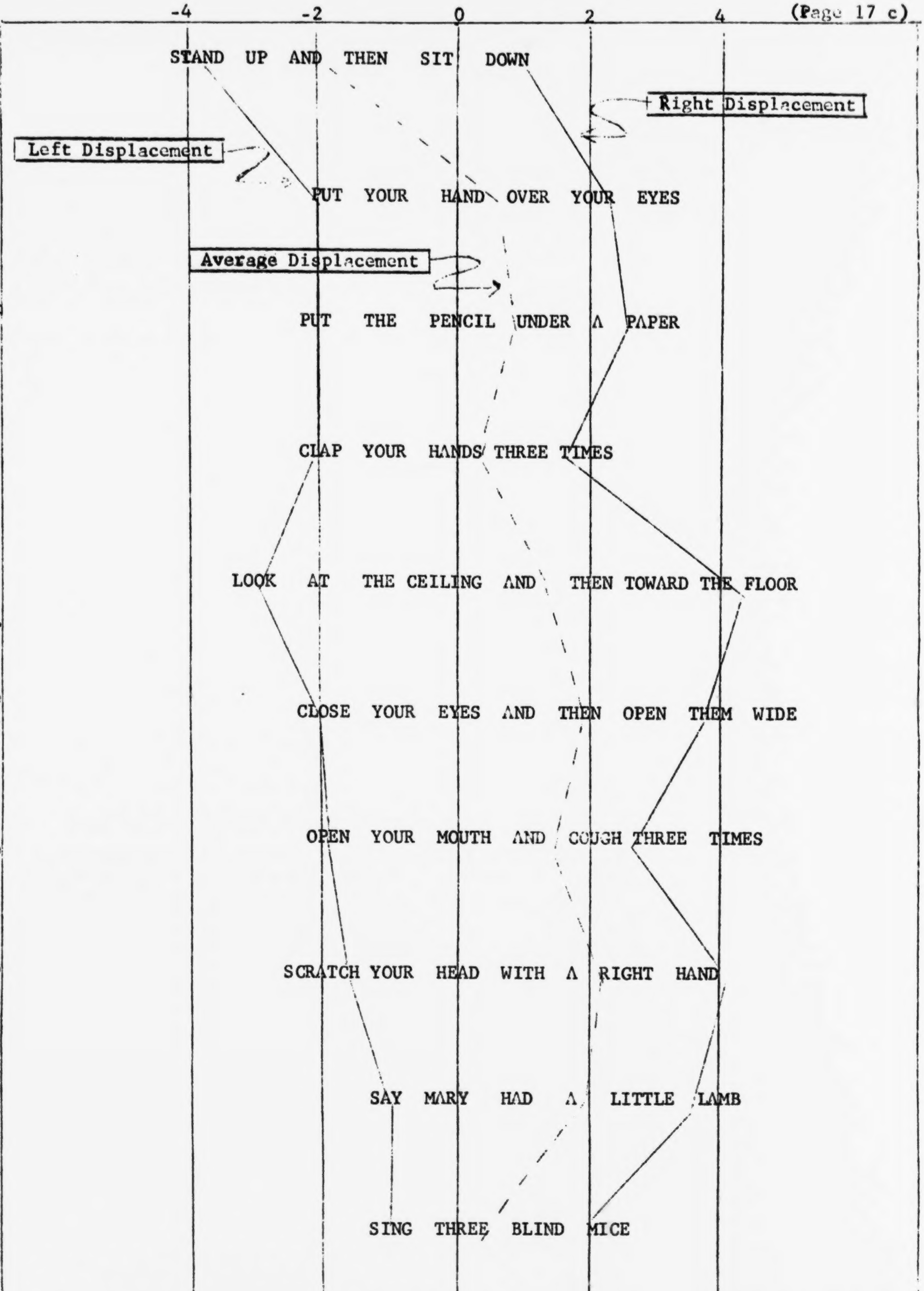


Figure 1 - Signed and Averaged Displacements for Sentence Strings

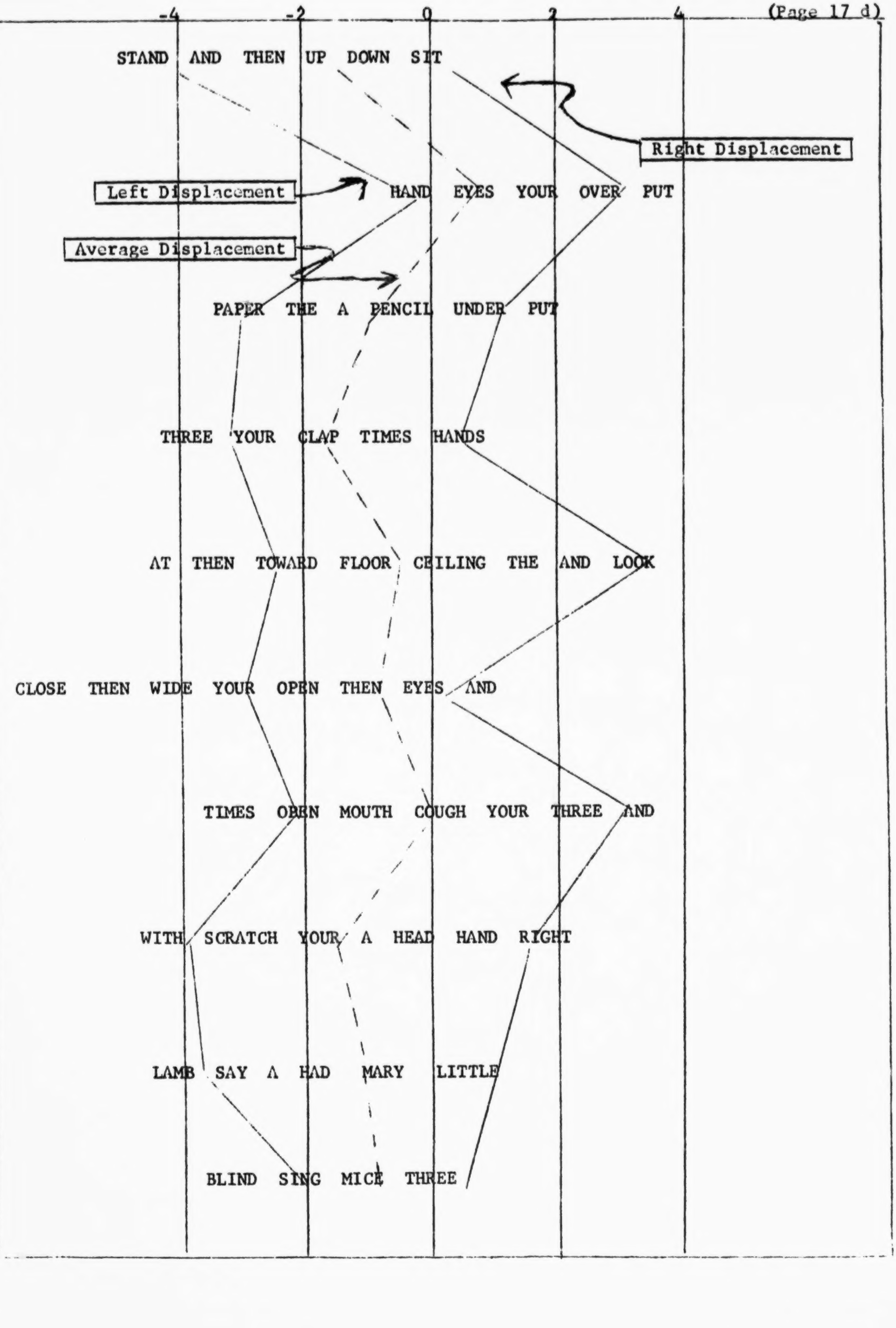
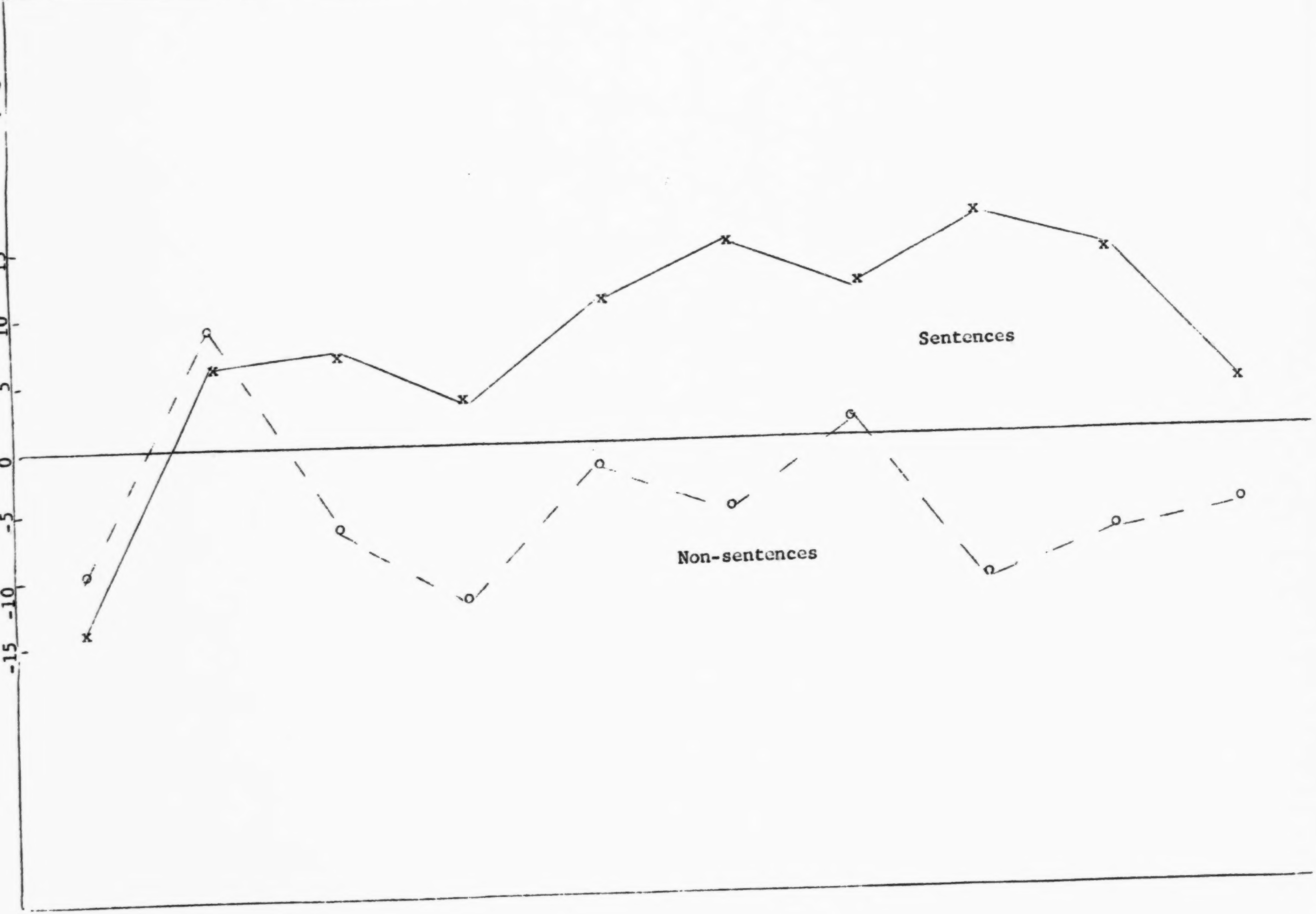


Figure 2 - Signed and Averaged Displacements for Random Strings

Figure 3 - Algebraic Displacements for the Combined Subject Groups



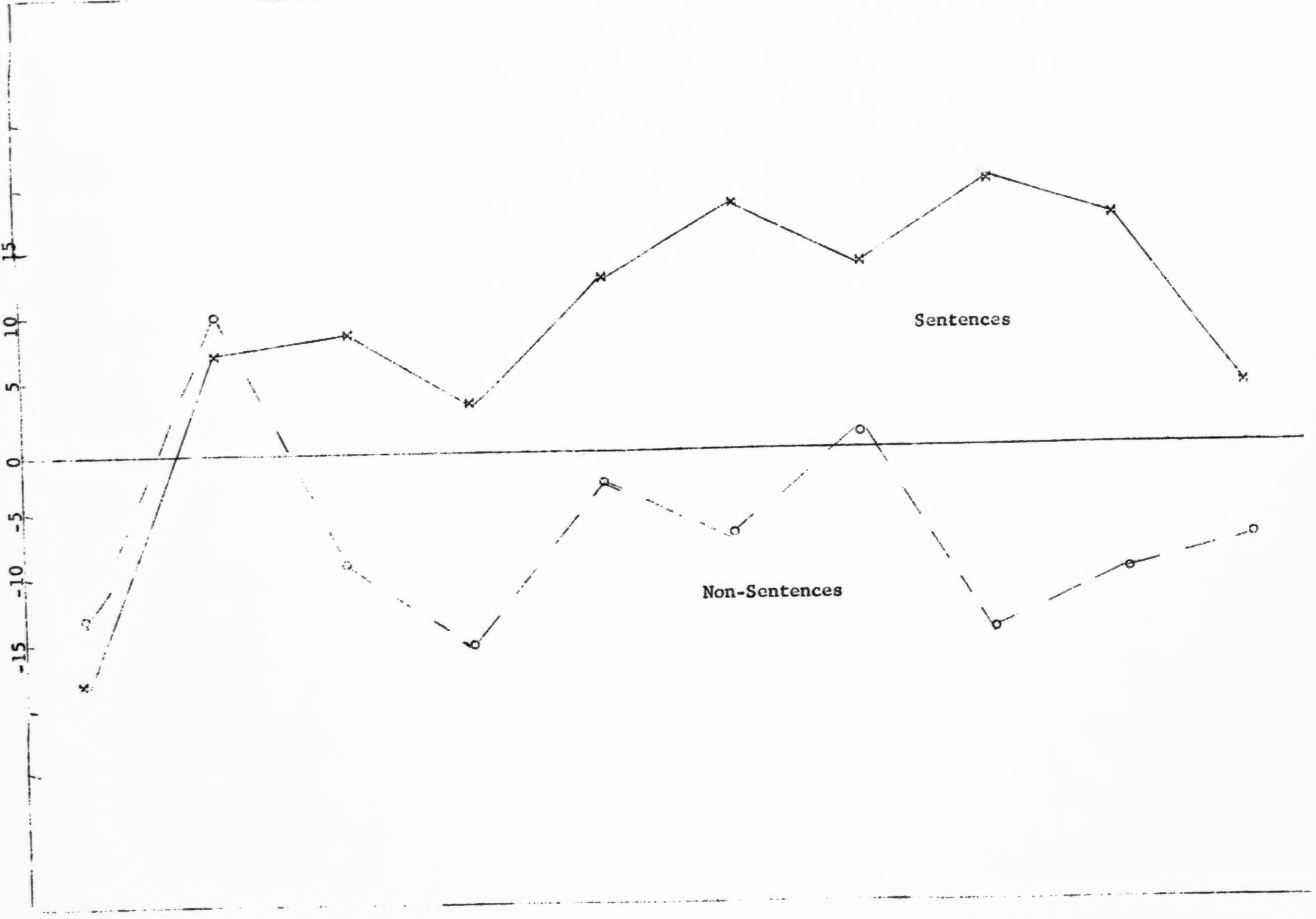
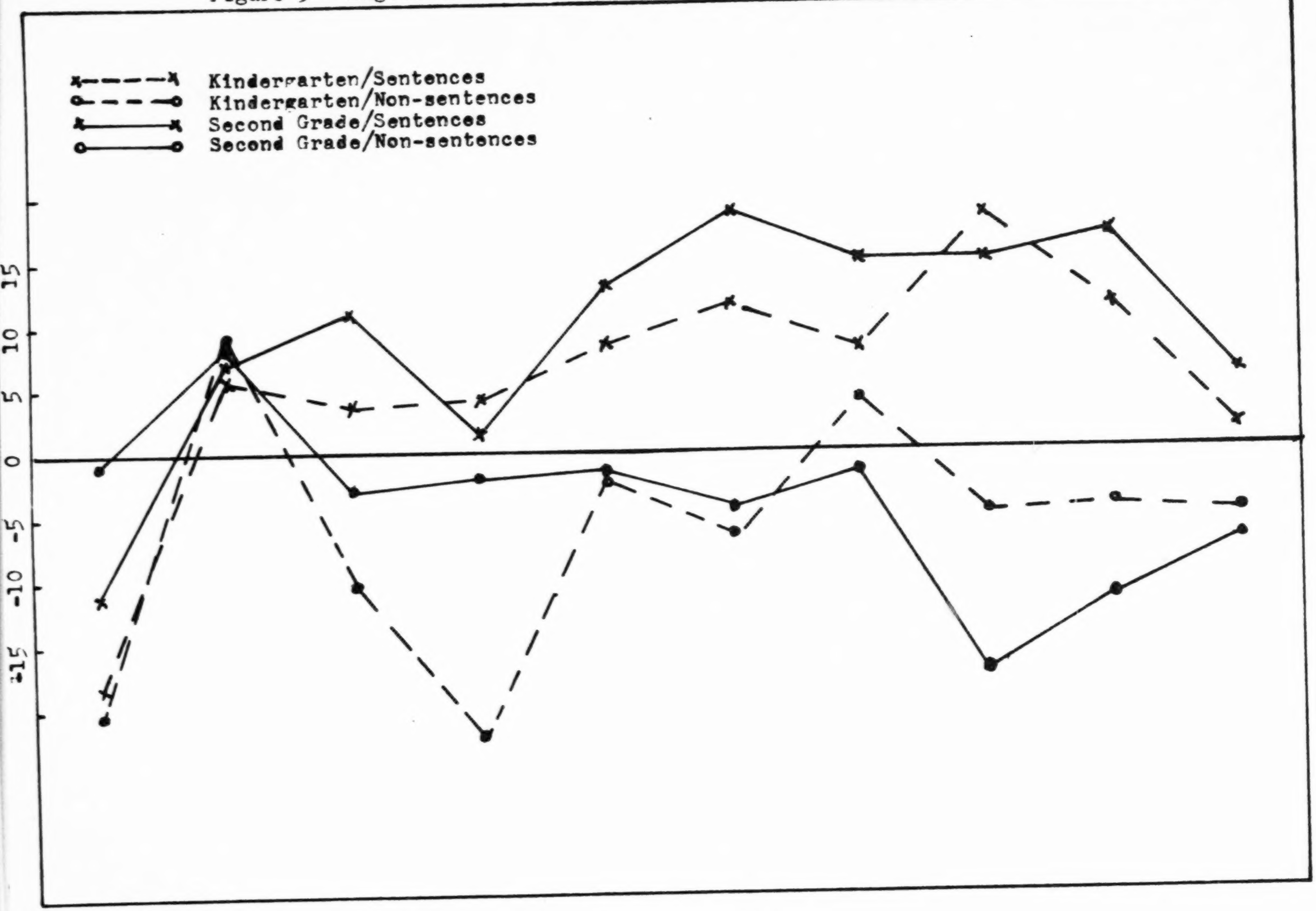


Figure 4 - Signed Absolute Displacements for the Combined Subject Groups

Figure 5 - Signed Absolute Displacements for All Classification Variables



in the direction of slightly greater displacement scores when direction of displacement (Figure 4) is preserved in summation. This lack of difference between the two forms of scoring is produced by a consistent tendency for Ss within a sentence or utterance type to displace the click locus in one direction.

The significant interaction between utterance composition and sentence in the algebraic analysis is very suggestive. The sentence/non-sentence contrasts of this experiment used exactly the same lexical stock. That is to say, in order to produce a non-sentence the words of a given sentence were simply randomly ordered. When one looks at the collapsed variance contribution of utterance composition, one is abstracting the effects of lexical content summed over its appearance in both sentences and non-sentences. Thus the (B)(C) interaction confirms the fact that the particular ordering of lexical content which produces a meaningful English utterance is peculiar in its effects upon displacement and that effect varies as a function of the particular lexical stock which is used to construct the sentence or non-sentence.

Figures 1 and 2 summarize the amount and direction of displacement for each of the utterances. These figures display each utterance in the scale units used for scoring, with the objective click locus centered on the zero displacement line. Both subject groups are combined in these displays. It is immediately apparent that there is a uniform tendency for the click location to be preposed from its objective location when non-sentences are used as stimuli and post-positioned when the utterances are well formed. This suggestive difference, however, must be qualified by the equally clear caveat that the condition of identical lexical locus for the click produced more clicks near the ends of non-sentences and at the beginnings of sentences. Thus the potential range of displacement scores is unequally limited in the two instances.

The reader will note from Figure 5, however, that there remains a distinct tendency for the kindergarten S̲s to prepose click locations for both sentences and non-sentences and for the second grade S̲s to post-position the click. The limitation of potential scoring range cited above is not an influence in this case, since the comparison involves both utterance types and is relative across grade level. As a post hoc check on this tendency a simple t̲ was calculated for the difference between the grade levels for the average differences between the sentence types. This t=4.34 and is significant beyond the 5% level. The test confirms what inspection of Figure 5 makes striking, that is, that although there is greater average displacement for sentences than for non-sentences at both grade levels, there is a greater tendency for kindergarten S̲s to preposition the click in non-sentences than there is for second grade S̲s.

It should be observed that the t̲ test just discussed is not a duplication of the (A)(B) interaction effect of the analyses of variance. The lack of significance in the ANOVAE merely indicates that the sentence and grade level functions are roughly parallel in their effects. The t̲ test by contrast declares that the differences in level of these functions, as displayed in Figure 5, are on the average significantly different.

<ins>Summary and Conclusions</ins>. In view of the relatively small N size employed in these experiments, the following general conclusions are subjectively strengthened in their prepotence as determiners of judgments of the locii of extraneous distortions in speech. For children, displacement of a tone-burst is controlled, at least in part, by the syntactic organization of lexical material, and by the nature of the lexical stock comprising that material. Greater magnitude of dis-

placement is found for sentences, lesser magnitude for non-sentences. For younger children there is a significant tendency to preposition the disturbed click location in both sentences and non-sentences. For both groups of children there is a tendency for click location to be preposed in non-sentential material and post-positioned in sentential material.

Procedural modifications involving the nature of the click distortion, its mode of presentation and the nature of the response task do not appear to delimit the generality of the observed findings. Locii identification for relatively young children is within their competence repetoire.

Children appear to be more sensitive to lexical modification within syntactic forms than do adults. And finally, as has been observed with adults, perceived click locii migrate towards the major constituent boundaries of sentential material.

Further research which would systematically vary the nature of the syntactic and lexical forms of the stimulus materials is clearly called for. The results of these experiments have indicated the fruitfulness and feasibility of such further research.

REFERENCES

Bellugi, U. The child's acquisition of grammar. Paper read at Eastern Association for Verbal Learning, 1967.

Chomsky, N. Syntactic Structures. The Hague, Mouton, 1957.

Fodor, J. A. and Bever, T. G. The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414-420.

Garrett, M. F. Syntactic Structures and Judgments of Auditory Events: A study of the Perception of Extraneous Noise in Sentences. Unpublished Ph.D. Dissertation, University of Illinois, Urbana, Illinois, 1965.

Garrett, M.F., Bever, T.G., and Fodor, J.A. The active use of grammar in speech perception. Journal of Perception and Psychophysics, 1966, 1, 30-32.

Ladefoged, P. and Broadbent, D. Perception of sequence in auditory events. Quarterly Journal of Experimental Psychology, 1960, 12, 162-170.

McCarthy, J. J. and Kirk, S. A. Illinois Test of Psycholinguistic Abilities: Experimental Edition. Urbana, University of Illinois, 1961.

Wells, R. Immediate constituents. Language, 1947, 23, 81-117.

Yngve, V. H. The depth hypothesis. IN: Jakobson (Ed.) Structure of Language and its Mathematic Aspects. Proceedings of Symposia in Applied Mathematics, 1961.

Acquisition and Transfer Differences Between Kindergarteners and Second-graders on Aurally and Visually Presented Paired-associates Using an A-B, A-C Design

Vernon C. Hall*

As Jensen and Rohwer (1965) point out, many conceptions about the learning process might be considerably different if psychologists had used subjects other than college sophomores in rote learning experiments. It is interesting that while many psychologists hypothesize developmental changes (be they maturationally or experientially based) in other areas, rote learning has until recently been both theoretically and empirically neglected. The present study was carried out to extend our knowledge about developmental differences in paired-associate learning. More specifically, this experimenter was concerned with the effects of mode of presentation on acquisition and interference with kindergarten and second-grade children.

In an earlier article Loomis and Hall (in press) using the aural mode of presentation employed the A-B, A-C design with kindergarten and second grade children. They found significant interference effects with the 8-year-olds but not the 5-year-olds. This was not only contrary to White's (1965 p. 214) prediction that 5-year-olds should exhibit significantly more negative interference but also fails to demonstrate a well-known phenomenon found many times with adults. This clearly demanded replication with a new sample. In addition, most studies unilizing the A-B, A-C paradigm have used visual presentation and it was felt that adding the visual mode here would considerably increase our knowledge about the generality of the Loomis and Hall findings. The Loomis and Hall study also used a double criterion (15 trials or 1 perfect trial) for its measurement of transfer effects and running everyone the same number of trials should again increase our confidence in the results. Finally, an additional pair was added to the lists (from four to five) since it was found that the younger children were easily able to handle this many.

---

* The author would like to thank Arlene Blake and Richard Kingsley for their valuable assistance.

With regard to acquisition, it has typically been reported that young children perform better with aural than visual presentation (i.e., McGeoch and Irion, 1952). Budoff and Quinlan (1964a) using 8 pairs of nouns and verbs from preprimers (some of which formed meaningful combinations; i.e., look - dog) found that 7 and 8-year-olds learned the lists significantly faster when presented aurally. They also used the double criterion of 1 perfect trial or 24 trials. Later, the same experimenters (1964b) replicated these findings using average and retarded readers with the same materials and procedures. Otto (1961) paired five common forms (i.e. triangle) with low association value trigrams which he presented to three grade levels (2, 4 and 6) of good, average and poor readers. Here the differences between procedures used in aural and visual presentation (he calls them reinforcements) modes are less pronounced (in both conditions the cvc was articulated but the visual condition included presentation of the three letters). Here, again, the second-graders (but not fourth and sixth) were better at the aural than visual presentation. There was an added problem, however, in that both serial order and scrambled order were used and the subject was run until he reached one perfect trial in each presentation type.

In no case have experimenters used kindergarten children or pictures rather than words in comparing visual with aural presentations. Since as White (1965) has pointed out, the age range between 5 and 7 has been identified by several theorists and much empirical data as being a transitional stage for learning processes, it is especially appropriate for any study interested in early developmental changes to include this age span. With regard to the pictures rather than words, not only is it impossible to use the latter with kindergarten children but it seems more appropriate to use pictures with second-graders since ability to read might well be a confounding factor when making a comparison with the aural presentation..

Method

Design and Sample. The basic design compared two age levels (kindergarten and second grade), two modes of presentation and two paradigms; (experimental and control). Subjects were 60 kindergarteners randomly chosen from approximately 120 kindergarteners and 60 second graders randomly chosen from about 125 students at Wetzel Road Elementary School, Liverpool, New York. The mean ages were 6.1 and 8.2. The subjects in each age group were randomly assigned to one of the four groups (aural or visual; experimental or control), which left 15 subjects in each cell.

Materials and Procedures. The lists used are shown in Table 1 and were nouns randomly selected from the 45 singular nouns used as stimuli in gathering the oral word association norms from young children by Palermo and Jenkins (1966). Outline drawings of the objects these words represented were then made, photographed and reproduced as 35mm slides. These slides were projected on a 9 inch by 9 inch screen (See Figure 1) by a Sawyer 707Q slide projector. The projector was operated automatically by two Hunter Timers. The first object was presented for 3 seconds by itself and then the two objects were shown together for three seconds. The inter-trial interval was 6 seconds.

Aural presentation was recorded on a Magnicord 1140 Tape recorder and presented to the children via Sharpe earphones. The stimulus word was first pronounced and then after a three second pause it was pronounced again with the response word. Both modes were in all other details identical. The actual lists were arranged in five random orders with the limitation that no single pair was to occur consecutively. This was violated on trial ten where the circular tray on the slide projector recycled and the last pair on trial nine was the same as the first pair on trial ten.

Prior to the experimental task, each subject was presented a two-pair list (knife-fork and ice cream-cake) to a criterion of one perfect trial to be sure the children understood the instructions. The instructions seen below were adapted from McCullers (1963).

> This is a game to see how well you can learn. If you try real hard we will give you some M and M candy afterwards. This is how it goes. When I turn on this slide projector (tape recorder) you will see (hear a word) an object. A few seconds later you will see (hear) the same object (the same word) with another object (word). Your job is to guess what the second object (word) will be, before you see (hear) it. You are to say the name of the object (word) out loud so I can hear you. Although the objects are all very common, I will tell you what we call them the first time through if we happen to call them different things. Here is an example of what you are going to see (hear). (1 trial with the two pair list) O.K. now we are ready to start (Present the two pair list until the subject gets one perfect trial and then clarify instructions if necessary.)

The underlined sentence was only used in the visual condition. Other condition differences are in parenthesis and should be self evident. The candy was used to keep the attention of the children and each subject received 4 M and M candies upon completion of the task.

All subjects repeated the first list until one perfect anticipation trial was performed and then immediately shifted to the second (within 30-60 seconds). The second list was presented to all subjects for 9 trials. Table 1 shows the lists used.

RESULTS

A 2 (experimental-control) x 2 (kindergarten-second grade) x 2 (audio-visual) analysis of variance was computed on trials to criterion on the first list. This analysis yielded a significant visual-audio effect ($F = 5.11$, 1/112 df, $p < .05$) and a significant experimental-control interaction with kindergarten and second grade. Examination of Table 2 which shows the appropriate means and standard deviations indicates that children reached criterion on the list significantly faster when it was visually presented. The interaction was caused by the kindergarten controls doing better than the experimental group while the second grade controls did not do as well as the experimental group. No other interactions or main effects were significant.

The second list was initially analyzed using a 2 (experimental-control) x 2 (aural-visual) x 2 (kindergarten-second grade) analysis of variance using a number of correct anticipations. This analysis yielded a significant experimental-control effect ($F = 8.69$, 1/112 df, $p < .01$) with no other main effects or interactions reaching significance. Subsequently two separate 2 (experimental-control) x 2 (aural-visual) analysis of variance were computed for each grade. The kindergarten analysis failed to yield any significant main effect or interaction. The second grade analysis yielded a significant experimental-control effect ($F = 7.98$, 1/56 df $p < .01$). The aural-visual main effect was not significant ($F = 3.55$, 1/56 df, $p < .05$). The appropriate means and standard deviations are shown in Table 3.

DE	AB	AC
doctor-chair	house-gun	house-bread
table-hammer	finger-woman	finger-sheep
bed-dog	needle-cheese	needle-girl
spider-man	lion-scissors	lion-boy
hand-lamp	baby-car	baby-mountain

Table I. Word List used.

X	KV	KΛ	SV	SΛ
E	7.2 4.57	7.8 3.95	7.4 3.45	8.1 4.67
C	7.2 2.94	11.2 5.52	5.7 2.49	6.4 3.49

Table II. Means and standard deviations for number of trials to criterion for first set.

	KV	KΛ	SV	SΛ
E	20.87 10.50	23.33 8.24	24.67 8.71	23.0 10.31
C	25.33 8.95	27.33 6.84	33.33 6.99	26.73 7.59

Table III. Means and s.d. of number correct on trials 2-9 on second set.

Discussion

Although it is, of course, impossible to exactly equate the two modes of presentation, the writer felt that by using pictures rather than printed words and earphones rather than free field he had come closer than previous attempts. The present writer also feels that in light of these findings the earlier generalization (McGeoch and Irion, 1952; Budoff and Quinlan, 1964a) that children learn faster aurally than visually must be modified. It seems possible that the conditions under which different modes excell may be quite specific with both subject and task parameters in need of further investigation.

The fact that 5-year-olds show no significant negative transfer in the A-B, A-C design has now been replicated three times (the present article, Loomis and Hall, in press, and an unpublished pilot study) with the audio mode. The addition of one more pair and the visual mode of presentation in the present study adds to our confidence that this is indeed a generalizable phenomenon. When we add this to the fact that other studies using this age group (i.e., Koppenaal, Krull and Katz, 1964; Jensen and Rohwer, 1965) have also found the 5-year-olds to perform differently from older children, the present author believes a case is made for asking psychologists interested in paired associate learning to give this age group more theoretical attention. Maybe the most pressing problem concerns the question of whether this is a maturational or experiential effect. The tendency so far, has been to try and account for it through experience (i.e., Koppenaal, Krull and Katz suggest amount of remote prior learning) and it is also the preference of the present experimenter. On the other hand, this still leaves unanswered the question of why most children acquire the correct amount of experience to behave like college students between the ages of 5 and 7. Since all children in these studies have been in some kind of school at time of testing and the present experiment was done in the latter half of kindergarten, after some formal instruction had begun, the effect of school may be less important than often represented.

References

Budoff, M. and Quinlan, S. Auditory and visual learning in primary grade children. Child Development, 1964, 35, 583-586.

Jensen, A. R. and Rohwer, W. D. Syntactical mediation of serial and paired-associate learning as a function of age. Child Development, 1965, 36, 601-608.

Koppenaal, R. J., Krull, A. and Katz, H. Age, interference, and forgetting. Journal of Experimental Child Psychology, 1964, 1, 360-375.

Loomis, D. and Hall, V. The effects of age and overlearning in an A-B, A-C design using high and low H value. Journal of Verbal Learning and Verbal Behavior, in press.

McCullers, J. C. An analysis of some factors underlying intralist associative transfer in paired-associate learning. Journal of Experimental Psychology, 1963, 65, 163-168.

McGeoch, J. A. and Irion, A. L. The psychology of human learning. New York: David McKay Company, Inc., 1952.

Otto, W. The acquisition and retention of paired associates by good, average, and poor readers. Journal of Educational Psychology, 1961, 52, 241-248.

Palermo, D. S. and Jenkins, J. J. Oral word association norms for children in grades one through four. January, 1966, 60.

White, S. H. Evidence for a hierarchial arrangement of learning processes. In L. P. Lipsitt and C. C. Spiker (Eds.) Advances in child development and behavior. New York: Academic Press, 1965.

Concept Identification Strategies

William J. Meyer and David Hultsch

It has been established that concept identification (CI) in adults is facilitated by procedures which decrease the memory load of the S (Bourne, Goldstein, and Link, 1964; Cahill, and Hovland, 1960; Dominowski, 1965; Pishkin, and Wolfgang, 1965). A recent investigation (Pishkin, Wolfgang and Rasmussen, 1967) extended these findings to adolescents, and found that the younger Ss benefited from the reduction of memory load to a greater extent than the older Ss.

Developmental studies of CI with young children have typically found increasingly better performance with increasing age (Meyer and Offenbach, 1962; Osler, and Kofsky, 1965, 1966). Osler and Kofsky (1965, 1966) have shown that one reason for these age differences is that younger children respond to the irrelevant aspects of the task to a greater extent than the older children. Osler and Kofsky (1966) also demonstrated that older Ss categorize stimulus information to a greater extent than younger Ss.

The purpose of the present investigation is to determine age differences in the effects of memory load on CI tasks at varying levels of complexity. The fact that older children seem to spontaneously categorize stimulus information to a greater extent than younger children (Osler and Kofsky, 1966) suggests that less of a memory load is placed upon the older Ss; that is, failure to categorize requires the S to recall the association between each separate stimulus and the correct response, while categorization of stimuli requires only that the S recall the association between a combination of stimuli and the correct response. Thus, it seemed plausible to expect that a reduction of the memory load required by the task would be of greater benefit to the younger than the older children. A corollary of this hypothesis is that reduction of the S's memory load would reduce responding to the irrelevant stimulus dimensions of the task for all age groups but to a greater degree for the younger as opposed to the older children.

Method

Subjects

The Ss were 108 boys and girls from kindergarten (K) and second grade (SG) classes at a suburban, Central New York school. The K sample contained 32 boys and 22 girls ranging in age from 5.6 to 6.7 years, with a mean of 5.9 years. The SG sample contained 30 boys and 24 girls ranging in age from 7.4 to 8.11 years, with a mean of 8.0 years. The 54 Ss within each age level were assigned at random to one of nine groups, resulting in six Ss per cell. The children are generally lower-middle class with generally average intellectual ability.

Stimuli

Concept identification tasks at three levels of stimulus complexity were constructed. Each level required the identification of one relevant concept, and complexity was varied by including either 0, 1, or 2 irrelevant concepts. In information theory nomenclature the problems contained either 1, 2, or 3 bits of information. Three bi-value concepts were combined to yield 12 possible problems shown in Table 1. Within each level of complexity, the relevant concept was assigned to each S at random.

The positions of the correct responses were sequenced according to criteria outlined by Fellows (1967). These sequences are designed to control responding on the basis of position perseveration, position alternation, win-stay, lose-shift, and win-shift, lose-stay strategies at a chance level. Thus, stimuli were sequenced so that within a block of trials (a) each stimulus appeared equally often, (b) correct responding involved selection of the right and left buttons equally often,

Table 1

Characteristics of the Experimental Problems

Characteristics[a]			
Bits of Information	Relevant Dimension	Irrelevant Dimension(s)	Number of Stimuli
1	Form	-	2
	Color	-	2
	Size	-	2
2	Form	Size	4
	Color	Form	4
	Size	Color	4
3	Form	Color Size	8
	Color	Form Size	8
	Size	Form Color	8

[a] Values of the concepts used were: form-circle, square; color-red, blue; size-one to two ratio.

and (c) shifting sides was rewarded equally as often as staying on the same side. Sequences 1 and 3 (121122211122 - 112221112212) from Fellows (1967) were combined to form a block of 24 correct response positions.

Design and Procedure

The basic design was a 3 x 3 x 2 fixed effects model varying three levels of complexity (1, 2, or 3 bits of information), three levels of memory load reduction (0, 1, or 2), and two grade levels (K or SG). The procedure consisted of three parts; pretesting, pretraining, and the experimental task. To prevent fatigue, the pretesting and pretraining were conducted on one day, and the CI task was administered the following day. All Ss were run individually.

During the pretesting each S was presented with several pictures, and was questioned by the E to determine ability to discriminate the forms, colors, and sizes used in the experiment. Subjects who were unable to make any one of these discriminations were not used in the experiment. One K and one SG S failed the pretesting.

Both the pretraining and CI tasks were presented visually on the same apparatus. The S was seated at a table in front of a milk glass screen, and the stimuli were rear projected on the screen terminating upon the child's response. A procedure of successive presentations was employed. Two buttons were placed on the table in front of the S, one to the right and one to the left of center. Activation of the correct response button operated an M&M dispenser (Davis Scientific Instruments, Model MMd-1). which delivered a candy to a box in view of the S. The equipment was automatically sequenced by means of a Gerbrands Program Timer (Model No. 1001).

For the memory load conditions in which 1 or 2 previous instances were available, a lighted display box was placed behind each of the response buttons. Following a correct response in these conditions, a print identical in content and size to the original projected stimulus was placed by the E in the box behind the appropriate response button. In conditions where one instance was available, the print remained in the display box until the S made another correct response, at which time it was removed and the new instance was placed behind the appropriate button. The procedure was the same for the conditions in which two previous instances were available, except that two prints were displayed at any one time. Following a third correct response, the earliest instance of the two displayed was removed. Only previously correct instances were used since other investigations (Pishkin, Wolfgang, and Rasmussen, 1967) have shown previous instances facilitate performance only if they are positive. For the 0 condition, no stimuli were given.

During the pretraining task, the Ss were shown pictures of a boy and a girl. The task was to press the left button whenever the girl appeared on the screen, and to press the right button for the boy. The E first showed the S which buttons to press, and explained the nature of the task. In the conditions where appropriate, the E also displayed the previously correct instances and explained their function. The Ss were encouraged to attend to these stimuli. Following the instructions, a series of the two pretraining stimuli were presented in blocks of eight trials with a correct response sequence of 12112221. The interstimulus interval was four seconds. Reinforcement was verbal, with the E saying "right" for correct responses and "wrong" for incorrect responses. The Ss were run to a criterion of 8 consecutive correct responses or 72 trials, whichever came first.

Subjects who failed to reach criterion within 72 trials were dropped from the experiment. The range of errors on this task was small (0 to 5), and only one SG S failed to reach criterion.

Prior to the CI task on the following day, the instructions were repeated and eight additional practice trials were given to all Ss using the pretraining stimuli. In the CI task, the S was required to respond without aid from the E. Only one problem was given to each S in which the task was to consistently place one value of the relevant dimension to the right, and the other value to the left by pressing the appropriate buttons. Successive presentations were used with an interstimulus interval of six seconds. Correct responses were reinforced with M & Ms, and no candy was delivered for an incorrect response. The Ss were run to a criterion of 10 consecutive correct responses or 144 trials, whichever came first.

Results

Total errors to criterion for each of the 18 groups are shown graphically in Figure 1 and Table 2. This table, as are all analyses, is based on data with a cut-off point of 50 rather than 144 trials. This cut-off point was used since the median number of trials to criterion was 14, and only 4 of the 23 Ss who went beyond 50 trials reached criterion before reaching 144 trials. Thus, analyses of the data based on a maximum of 50 trials accurately reflects the results of the experiment.

Analysis of variance on the error scores yielded a significant bits effect ($F = 5.88$, d.f. = 2.90; $p < .005$) which accounts for 8% of the total variance. The analysis also indicated a significant memory load effect and accounts for 9% of

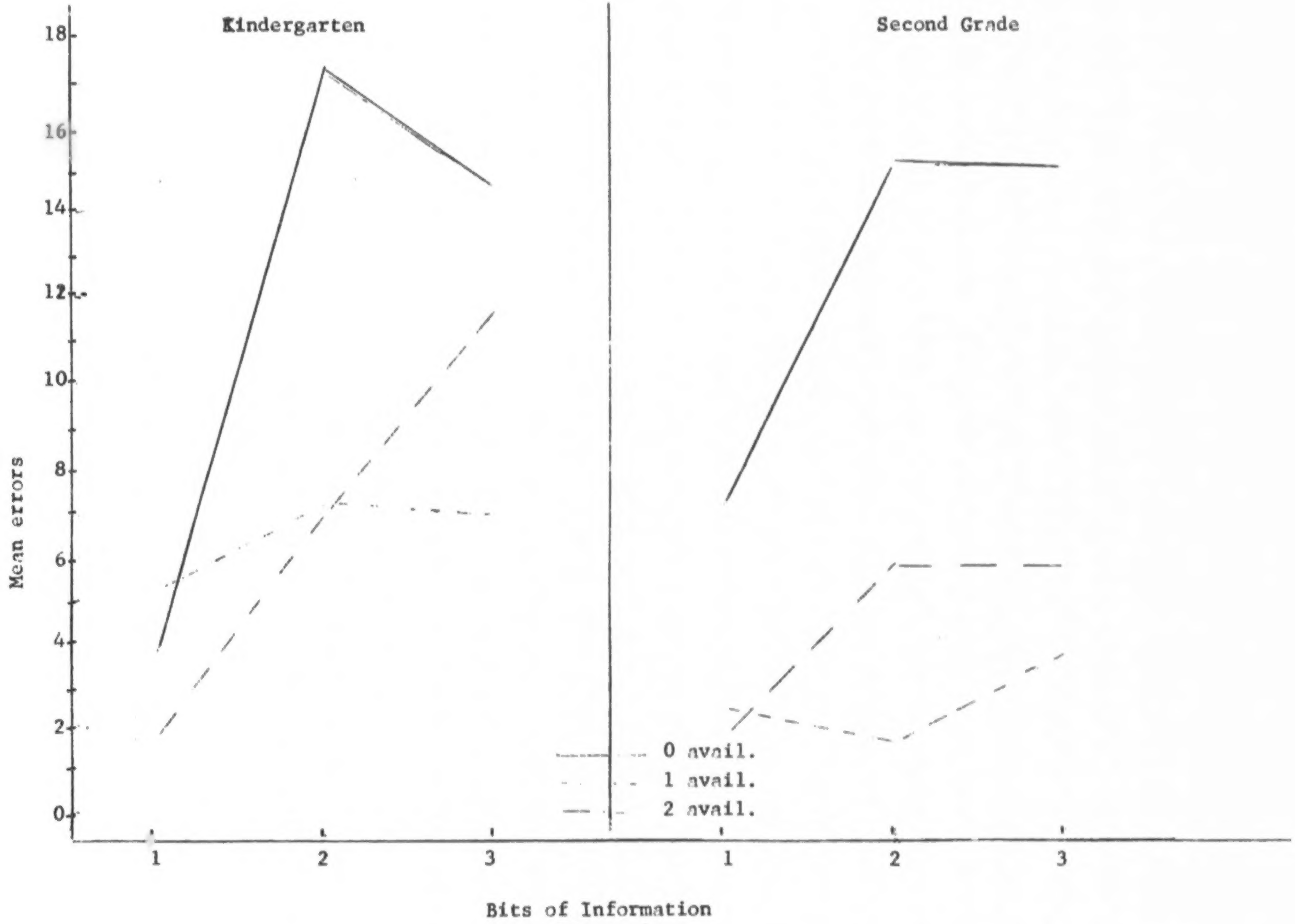


Figure 1
Mean errors for kindergarten and second grade groups at 1, 2, and 3 bits of complexity and 0, 1, and 2 available instances.

the total variance ($F = 6.74$; d.f. = 2,90, $p < .005$). No other main effect or interaction was significant.

Analysis of the simple effects for bits of information, and number of available instances was done to locate the sources of variation contributing to the significant main effects. The analysis for bits of information indicated a significant bits effect at 0 available instances ($F = 5.16$; d.f.= 2, 90; $p < .01$), but not at 1 or 2 available instances (see figure 2). Further comparisons with the 0 available instances indicated greater numbers of errors were made at the 2 and 3 bit levels, than at the 1 bit level ($F = 10.14$; d.f.= 1,90; $p < .005$).

Analysis of the simple effects for memory load yielded a significant effect at the 2bit level ($F = 6.18$; d.f.= 2,90; $p < .005$), but not at the 1 or 3 bit levels, although the comparison at the 3 bit level was significant at the .10 level (see figure 3). Further comparisons within the 2 bit level indicated fewer errors were made with 1 and 2 available instances than when no instances were available ($F = 12.02$; d.f.= 1,90; $p < .005$).

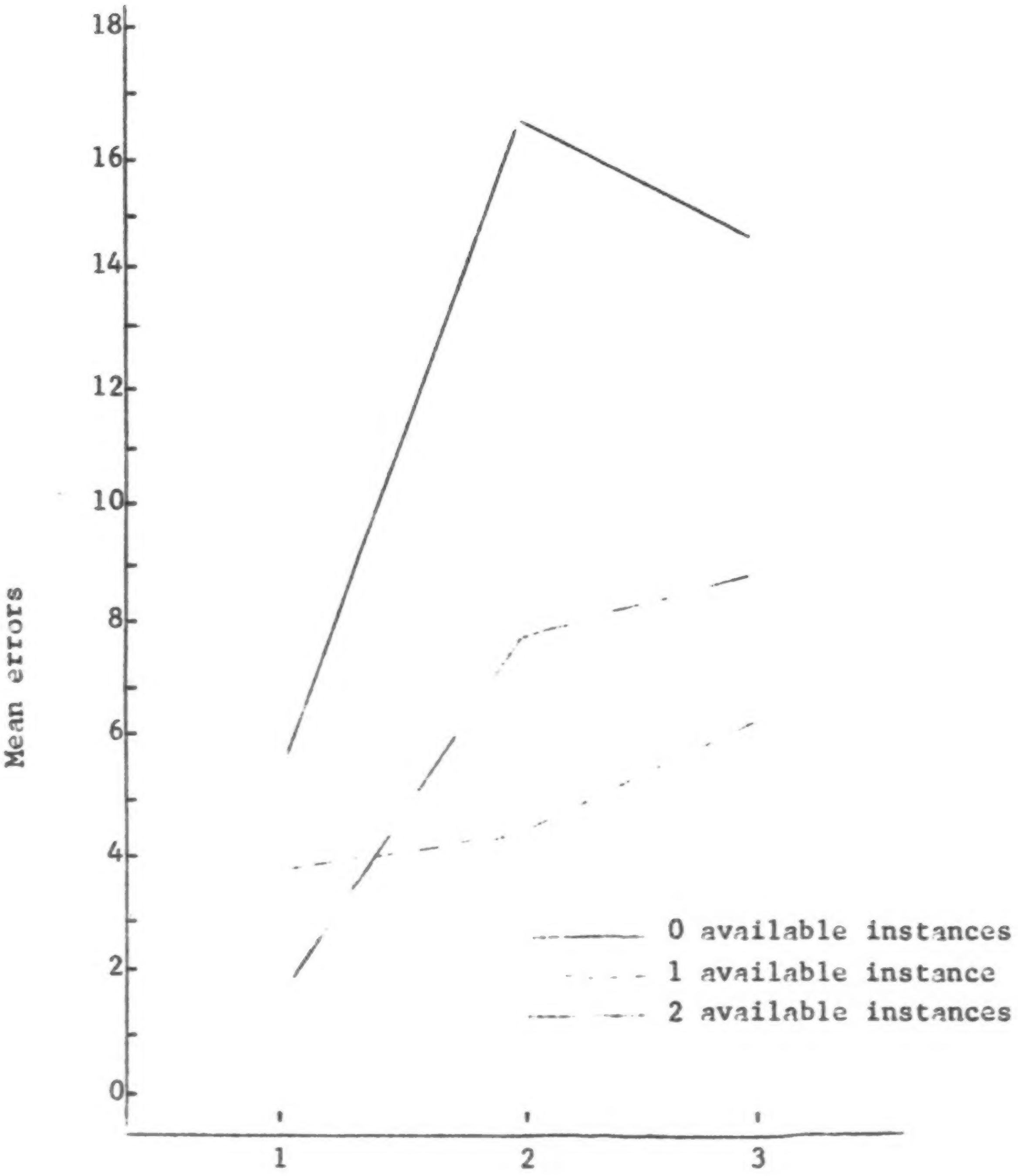
Since both bits of information, and number of available instances represent equal intervals of quantitative variables, trend analysis was performed on these factors by analyzing the sums of squares of the main effects into linear and quadratic components. It was found that the linear component accounted for 91% of the total sums of squares due to bits of information ($F = 10.67$; d.f.= 1,90; $p < .005$), while the quadratic component accounted for the remaining 9% of the variance which is not statistically significant. The linear component for the number of available instances was also significant ($F = 9.73$; d.f.= 1,90; $p < .005$) and accounts for 72% of the quadratic component accounted for the remaining 28% and was not significant.

A chi-square analysis of trials to criterion was carried out according to a technique described by Bresnahan and Shapiro (1966) in which an overall contingency

Table 2 - Means and variances of errors

			Kindergarten			Second		
			# available			# available		
			0	1	2	0	1	2
BITS	1	$\overline{X}$	3.8	5.2	1.5	7.2	2.7	1.8
		SD	6.1	9.8	1.2	10.0	4.2	1.7
	2	$\overline{X}$	17.2	7.2	7.0	15.2	1.5	5.8
		SD	11.1	9.6	9.4	12.3	2.3	8.7
	3	$\overline{X}$	14.3	12.0	11.8	15.0	3.8	5.8
		SD	10.7	13.8	9.5	12.2	3.2	6.4

F max = 127.73
F max, .01,5 = 88.0



Bits of Information

Figure 2

Mean errors for kindergarten and second grade groups combined at 1, 2, and 3 bits of information and 0, 1, and 2 available previous instances.

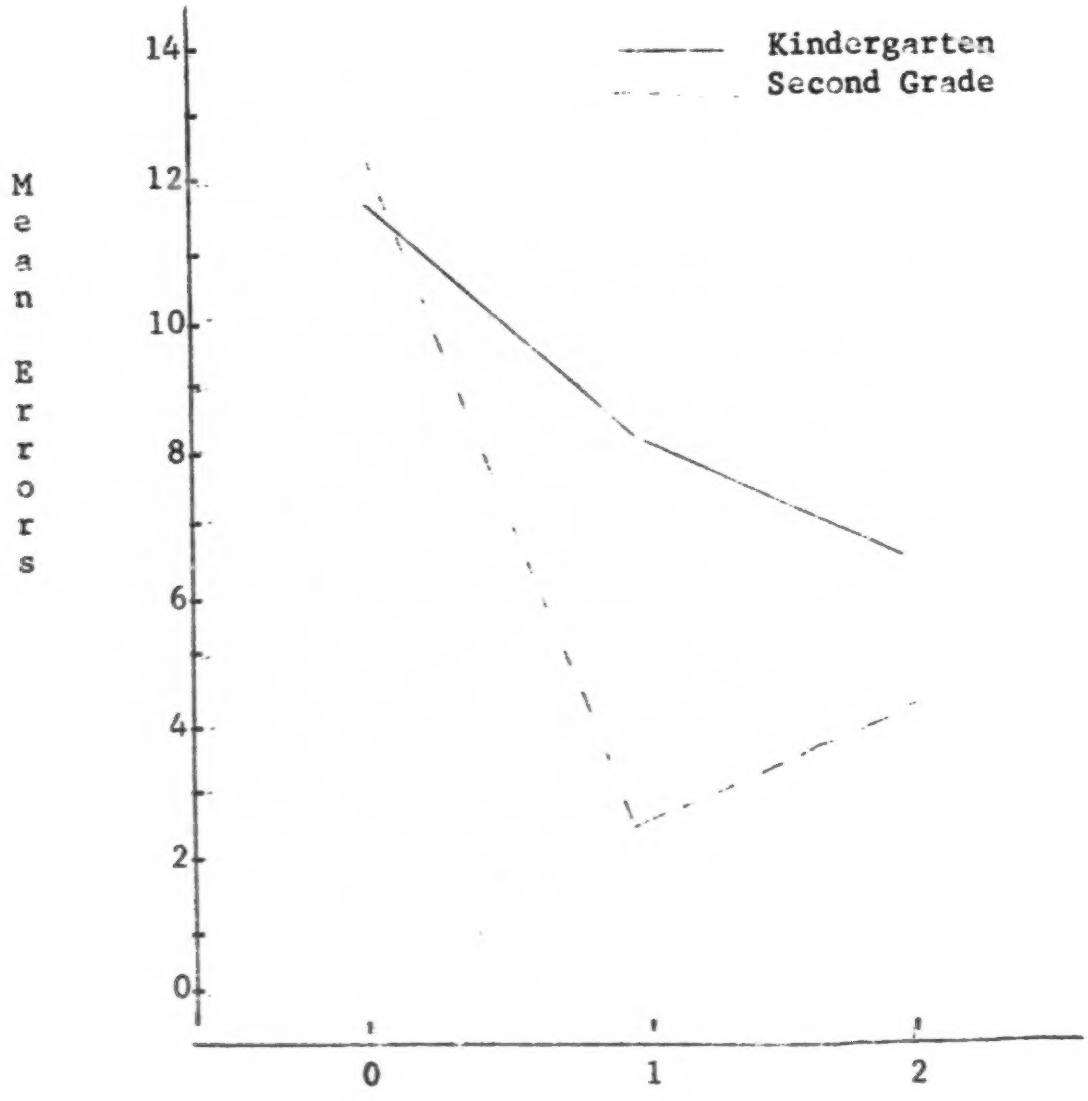


Figure 3

Mean errors for Kindergarten and Second Grade groups at 0, 1, and 2 available previous instances.

table with more than 1 *d.f.* may be partitioned into orthogonal components, each distributed as chi-square with 1 *d.f.* This permits precise analysis of the source of nonindependence which may be found in the overall contingency table. This analysis was carried out for several reasons. First, it was necessary to analyze the trials to criterion measure by non-parametric measures since a number of *S*s did not reach criterion. That is, there was no way to determine the number of additional trials these *S* would have required to reach criterion, and thus, these *S*s constituted only a category. Second, a non-parametric analysis was considered informative since the variances of the error scores were unequal ($F_{MAX} = 127.73$; *d.f.*= 5; $p < .01$). Since trials to criterion correlated .97 with errors, analysis on the former measure should provide a replication of the error analysis of variance results. Thus, while the *F* test has been shown to be robust with respect to the assumption of equal variances, replication of the parametric analysis by the non-parametric would increase confidence in the results of the experiment. Finally, the trials to criterion measure was used since it permitted comparisons of criterion *S*s with non-criterion *S*s, as well as further comparisons within the criterion group alone. Although these comparisons were possible analysis of variance of error scores, the procedure would have resulted in unequal numbers of *S*s within the cells, which, with unequal variances, seriously violates the assumptions underlying the *F* test.

Categories were formed by dividing the criterion *S*s into quarters on the basis of the number of trials required to reach criterion. Subjects not reaching criterion formed a fifth category. The experimental groups conform to the requirements for the analysis (Bresnahan & Shapiro, 1966). Table 2 shows the overall contingency table, analysis of which indicated significant non-independence

(X^2 = 62.20, d.f. = 44; $p < .05$). The overall contingency table was partitioned into 10 orthogonal components in an attempt to locate the sources of the non-independence indicated by the overall analysis. It was decided a priori to calculate these orthogonal components for criterion versus non-criterion Ss, and for a median split on trials to criterion within the criterion group alone. The results show that proportionally more Ss failed to reach criterion at 3 bits than at 1 bit, and that at 2 and 3 bits, proportionally more Ss reached criterion when 1 and 2 previous instances were available than when none were available. None of the other partitions yielded significant chi-square values. Thus, this analysis lends support to the results of the error analysis of variance, except that it indicates a significant effect due to available instances at 3 bits, while this effect only approached a low level of significance in the error analysis.

The partitions for the comparison of the criterion Ss alone, divided into those above and below the median on the trials to criterion measure, revealed that proportionally more Ss were above the median on the trials to criterion measure at 3 bits than at 1 bit. No other comparisons were significant. The effects found previously for the number of available instances were not found in this analysis, suggesting that these effects were due entirely to Ss failing to reach criterion.

Information theory suggests another data analysis strategy. Essentially, the question asked is how much of the information contained in the stimuli is actually transmitted to the child. For example, in information theory terms, a problem may contain one bit of information (two equi-probable events). In the case of perfect transmission, the S will consistently make one response for one event, and a second response for the other event. No information is transmitted if the S responds randomly to the two events. Thus, transmitted information may be thought

Table 3

Overall Contingency Table Tabulating Trials

To Criterion For All Treatment Groups

Experimental Group			Trials				
			10	11-12	13-16	17-43	Non-Criterion
1 Bit	0	K	3	1	0	2	0
		SG	0	2	1	2	1
	1+2	K	3	6	2	0	1
		SG	4	3	3	2	0
2 Bits	0	K	0	0	2	0	4
		SG	2	0	0	1	3
	1+2	K	2	1	3	4	2
		SG	2	4	3	2	1
3 Bits	0	K	0	1	0	2	3
		SG	0	1	1	1	3
	1+2	K	3	0	1	4	4
		SG	1	1	5	4	1

For this Table and other Tables of the chi-square analysis the following symbols will be used: c= Criterion Ss; NC= Noncriterion Ss; 0= No available Instances; 1+2= The sum of the 1 and 2 available Instance conditions; K= Kindergarten; SG= Second Grade.

of as the amount of reduction of uncertainty produced by the *S*'s response to the stimulus events. In this study, the analysis was concerned with the amount of information transmitted from the relevant and irrelevant stimulus dimensions and permits a direct assessment of the amount of systematic responding to the relevant and irrelevant dimensions. The use of error measures does not permit direct assessment because they may be caused by wholly random responding as well as by responding to irrelevant cues.

Details for computing the information transmission index are described by Attneave (1959, pp. 42-67). Essentially the computations are concerned with a comparison of the maximum possible uncertainty with actual uncertainty. Maximum possible uncertainty occurs when the subject is responding on a purely random basis. The difference between maximum uncertainty and actual uncertainty is the amount of reduction of uncertainty provided by the subjects responses.

Table 4 and figure 4 show the mean amount of information transmitted from the relevant and irrelevant dimensions for both the criterion and non-criterion *S*s of the various experimental groups. The values represent transmission scores between a maximum of 1.00 (perfect transmission), and a minimum of .00 (random responding). The significance of these values (difference from zero) was calculated by means of the chi-square formula: $X^2 = 1.3863\ N\ (T - \underline{d}.\underline{f}./\ 1.3863\ N$, where N equals the total number of responses, T equals the mean transmission score, and *d*.*f*. equals the degrees of freedom (Attneave, 1959, pp. 63-66).

The findings suggest that criterion *S*s discarded the irrelevant cues rapidly, and responded mainly to the relevant cues. For the criterion *S*s, 96% of the transmission was accounted for by the relevant cues, while only 4% was accounted for by the irrelevant dimensions. On the other hand, 20% of the non-criterion *S*s' trans-

Table 4

Mean Amount of Information Transmitted
From Relevant and Irrelevant Stimulus Attributes

Condition	Criterion Ss			Non-Criterion Ss		
	n	relevant	Irrelevant[1]	n	relevant	Irrelevant
Kgn 1 Bit 0 Available	6	.635^{d}	-	0	-	-
Kgn 1 Bit 1 Available	5	.678^{d}	-	1	.000	-
Kgn 1 Bit 2 Available	6	.562^{d}	-	0	-	-
Kgn 2 Bits 0 Available	2	.286^{c}	.000	4	.000	.040
Kgn 2 Bits 1 Available	5	.482^{d}	.050^{b}	1	.000	.000
Kgn 2 Bits 2 Available	5	.482^{d}	.079^{c}	1	.000	.000
Kgn 3 Bits 0 Available	3	.199^{d}	.012	3	.029^{b}	.030^{b}
Kgn 3 Bits 1 Available	4	.648^{d}	.041	2	.007	.045^{b}
Kgn 3 Bits 2 Available	4	.325	.022	2	.008	.001
2nd 1 Bit 0 Available	5	.445^{d}	-	1	.000	-
2nd 1 Bit 1 Available	6	.569^{d}	-	0	-	-
2nd 1 Bit 2 Available	6	.570^{d}	-	0	-	-
2nd 2 Bits 0 Available	3	.641^{d}	.008	3	.000	.020^{a}
2nd 2 Bits 1 Available	6	.680^{d}	.000	0	-	-
2nd 2 Bits 2 Available	5	.524^{e}	.008	1	.015	.082^{b}

Table 4
(continued)

Condition	Criterion Ss			Non-Criterion		
	n	relevant	Irrelevant	n	relevant	Irrelevant
2nd 3 Bit 0 Available	3	.378[d]	.036	3	.009	.021[a]
2nd 3 Bit 1 Available	6	.367[d]	.089[d]	0	-	-
2nd 3 Bit 2 Available	5	.413[d]	.019	1	.044	.193[d]

[1] The irrelevant dimension which transmitted the greatest amount of information for each S was used in computing the mean.

a $p < .05$

b $p < .025$

c $p < .005$

d $p < .001$

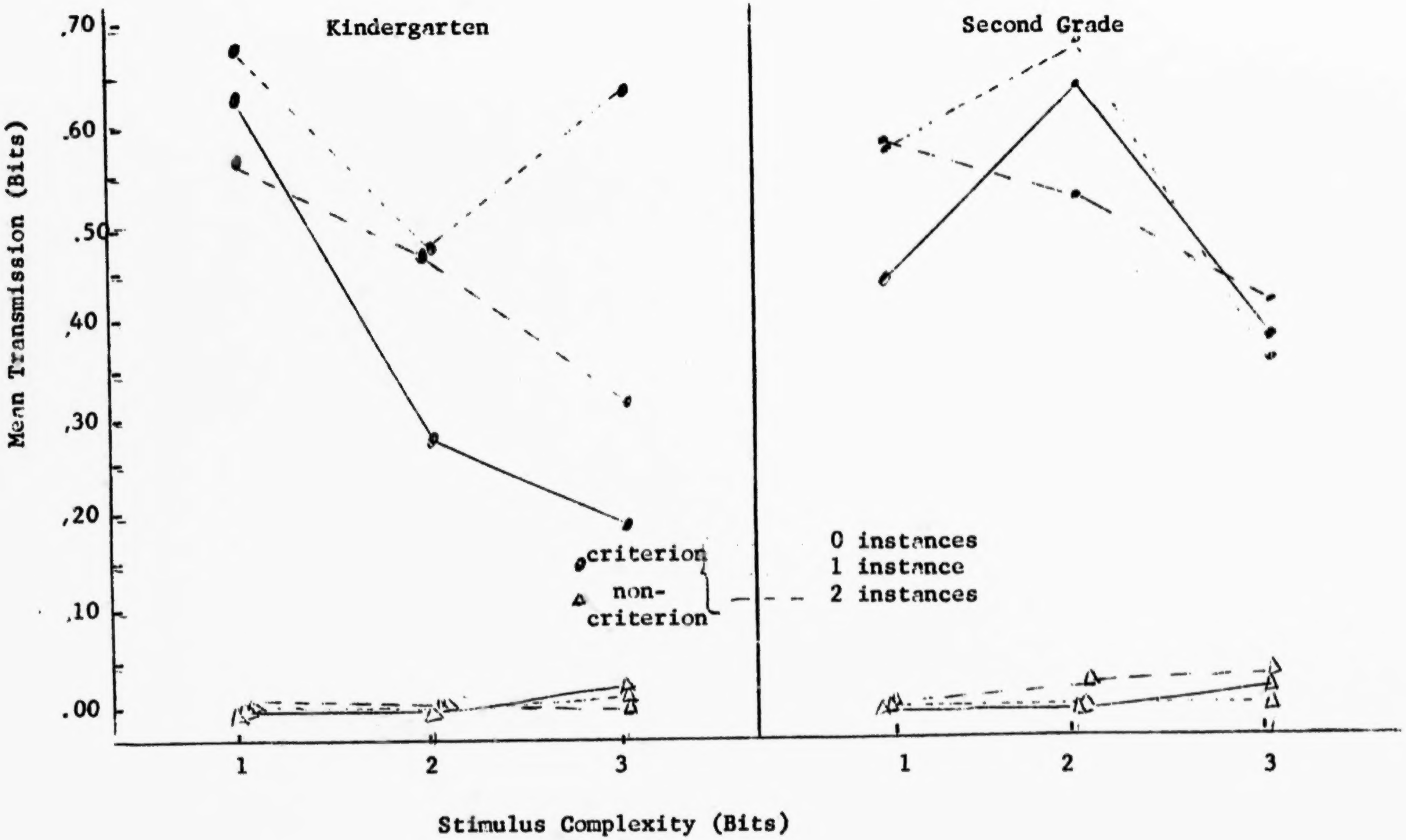


Figure 4

Mean Transmission (Bits) on relevant dimension for criterion and non-criterion Kindergarten and Second Grade groups as a function of stimulus complexity and number of available instances.

mission was accounted for by the relevant dimension, while 80% was accounted for by response to the irrelevant cues. Thus, it appears that response to irrelevant cues was a major factor in the failure of Ss to reach criterion.

Discussion

A major purpose of this study was to examine the interaction of age and task complexity where complexity was defined in terms of stimulus uncertainty. Although the stimulus complexity effect was statistically significant, the interaction failed to attain an acceptable level of statistical significance. Examination of figure 5 suggests, however, that the younger children were more adversely affected by the increase in stimulus uncertainty than the older children. The major reason, as suggested by the data, for the nonsignificant interaction effect is apparently the extreme variability of the children's performance. Inspection of the individual performance scores suggests that many of the children entered the learning task with the correct hypothesis already formed and, therefore, required relatively few learning trials. Other children either had no hypothesis or contrary hypothesis resulting in many more trials to achieve criterion. With only six Ss in a cell, the effect of the variability was to eradicate statistical significance.

It should be noted, however, that the general direction of our age x complexity interaction is entirely consistent with that reported by Osler and Kofsky (1965). Considering the two studies together, there is support for the conclusion that concept learning is adversely influenced by higher levels of stimulus uncertainty. The conclusion is further warranted that younger children are more adversely influenced by higher levels of stimulus complexity than older children.

One reasonable interpretation offered by Osler and Kofsky (1965) for the age x complexity interaction is that younger, as opposed to older children rely more

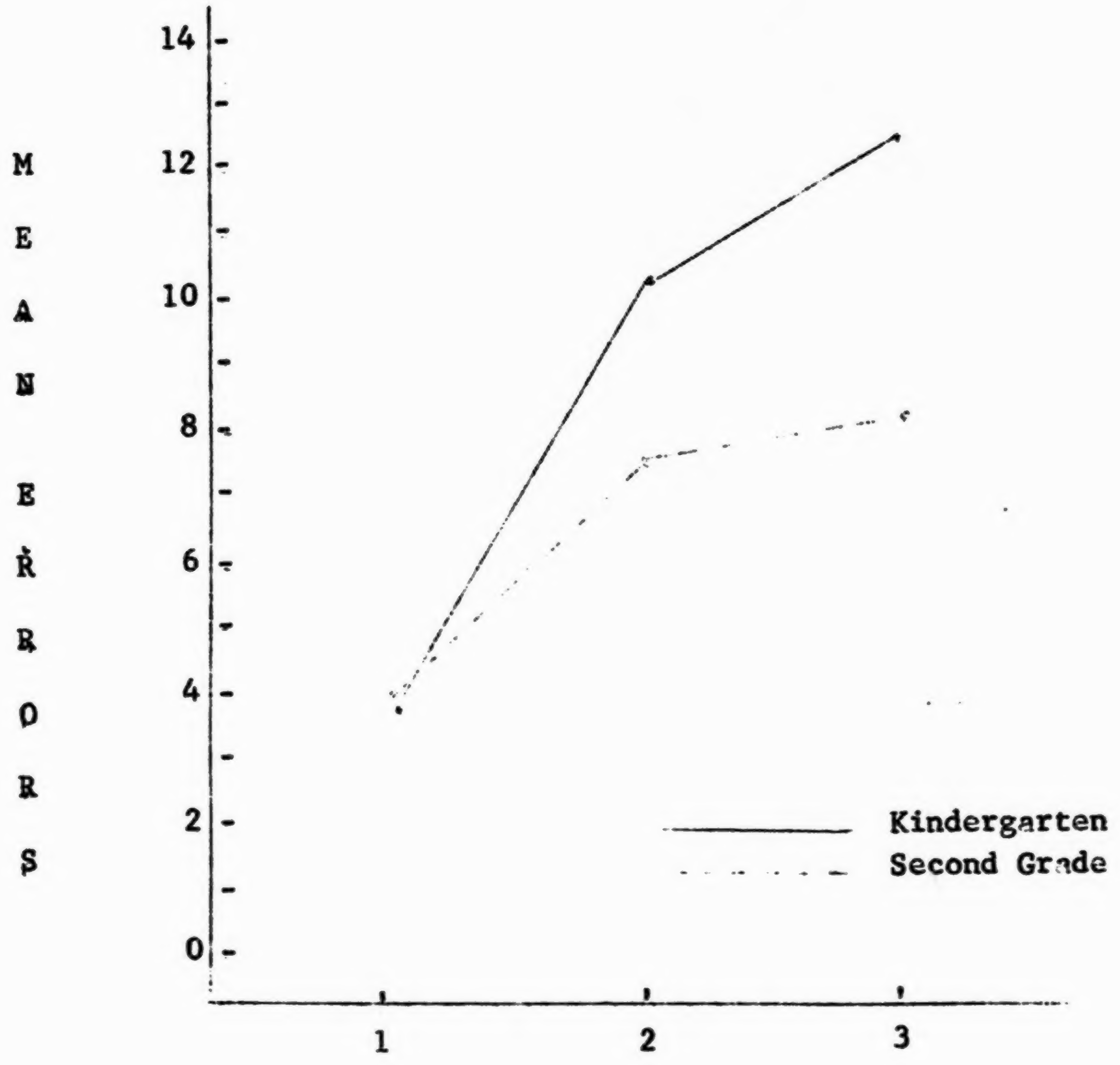


Figure 5

Mean errors for Kindergarten and Second Grade groups at 1,2, and 3 bits of information complexity.

on associating specific stimuli with specific responses. Older children, conversely, are probably dimensionalizing the task placing less of a load on their memory capacities. This hypothesis was examined in this study by providing the children with varying amounts of mamory aids. We anticipate an age x memory aid effect but, again, it failed to attain an acceptable level of significance. The effect attributable to the memory aid variable was statistically significant with the zero availability group performing less well than the one and two availability group. The interaction of stimulus complexity x memory aid was also not statistically significant.

Inspection of figures 1 and 3 suggest that the general trends are not consistent with the age x memory aid hypothesis. Figure 1 indicates that the second grade group improved more with the memory aid than the kindergarten group. Similarly, figure 3, which compares age groups for pooled complexity levels, indicates greater improvement for the older children. In terms of the original hypothesis these results suggest that the memory aid is more effective when the children dimensionalize the stimuli. Thus, in the task used in this study, the children dimensionalizing the stimuli, as opposed to the S-R learners, maintain their memory load advantage. Indeed, it can be argued that the memory aids provide a greater percentage reduction in memory load, because of dimensionalizing for the older, as opposed to, the younger children. This revised hypothesis will be examined in our next study.

A finding of particular interest from the information analysis concerns the performance of the non-criterion Ss. Although the number of non-criterion children in lack of the cells is small, it should be noted that they uniformly perseverated in their responses to the irrelevant dimensions. Another possibility, observable from this analysis, was for the children to respond randomly. This tendency toward perseveration was noted by Osler and Kofsky (1965) who also found a greater tendency among their younger Ss (four-year-olds) for that type of behavior. (Our data do not permit an age comparison because of too few Ss in the cells.) One might speculate that the children showing the perseverative behavior were satisfied with the approximately 50 1/0 reinforcement they received by continuously responding to the irrelevant dimension.

In view of the subject variability encountered in this study, we plan to change the design. Clearly, one feature of the design must include a greater number of Ss in each cell. Approximately 20 subjects per cell should be adequate. Secondly, we plan to include only two levels of stimulus complexity; one bit vs. 3 bits. Thirdly, the memory aid variable will include zero available instances as opposed to one available instance. Essentially the revised design is a 2 x 2 x 2 with two levels of stimulus complexity, two levels of memory aid, and two age levels.

Another modification derived from the present study involves the procedure. Instead of using only one relevant dimension, we plan to counterbalance, within groups, the relevant dimension. In this way it should be possible to counteract response preferences and it also makes it possible to partial out the effects of stimulus preferences, if they should occur.

Finally, we are planning to use "culturally deprived" children in our sample and compare their performance with another sample of middle-class children.

References

Attneave, F. *Applications of information theory to Psychology.* New York: Holt Rinehart and Winston 1959.

Bourne, L.E.,Jr., Goldstein, S., and Link, W.E. Concept learning as a function of availability of previously presented information. *Journal of Experimental Psychology*, 1964, 67, 439-448.

Bresnahan, J.L., and Shapiro, M.M. A general equation and technique for the exact Partitioning of chi-square contingency tables. *Psychological Bulletin*, 1966 66, 252-262.

Cahill, H.E., and Hovland, C.I. The role of memory in the acquisition of concepts. *Journal of Experimental Psychology*, 1960, 59, 137-144.

Dominowski, R.L. The role of memory in concept learning. *Psychology Bulletin*, 1965, 63, 271-280.

Fellows, B.J. Change stimulus sequences for discrimination tasks. *Psychological Bulletin*, 1967, 67, 87-92.

Meyer, W.J., and Offenbach, S.I. Effectiveness of reward and punishment as a function of task complexity. *Journal of Comparative and Physiological Psychology*, 1962, 55, 532-534.

Osler, S.F., and Kofsky, E. Stimulus Uncertainty as a variable in the development of conceptual ability. *Journal of Experimental Child Psychology*, 1965, 2, 264-279.

Osler, S.F., and Kofsky, E. Structure and strategy in concept learning. *Journal of Experimental Child Psychology*, 1966, 4, 198-209.

Pishkin, V., and Wolfgang, A. Number and type of available instances in concept learning. *Journal of Experimental Psychology*, 1965, 69, 5-8.

Pishkin, V., and Wolfgang, A,, and Rasmussen, E. Age, sex, amount, and type of memory information in concept learning. *Journal of Experimental Psychology*, 1967, 73, 121-124.

Abstract

Fear and Attachment in Young Children

J. Conrad Schwarz

Syracuse University

Children 3½ to 5½ years old were exposed to a fear stimulus either in the presence of their mothers or in the presence of an adult female stranger in a test of a hypothesis derived from attachment theory. Ratings were made of facial expression, movement about the room, and visual orientation over a 5 minute period. The results failed to support the prediction of less fear in the presence of the mother. At the appearance of the fear stimulus there was no change in facial expression for Ss in the stranger condition, whereas, Ss in the presence of the mother changed facial expressions in the unpleasant direction. Other comparisons, although nonsignificant, were also contrary to prediction. It was suggested that the presence of the stranger inhibited motility and emotional communication.

Fear and Attachment in Young Children[1]

J. Conrad Schwarz*

Syracuse University

The purpose of this study was to test the effect of the presence of an attached individual (the mother), as compared with the presence of an adult female of short acquaintance, upon the apprehensiveness of children when exposed to novel stimuli. The concept of attachment used in this context is synonomous with "specific attachment" as used by Schaffer and Emerson (1964). It is manifested behaviorally by a preference for the proximity of a specific individual and by a protest reaction to separation from the individual. It is also related to the concepts of attachment used by Cairns (1966) and of primary socialization bond used by Scott (1963).

Both Cairns (1966) and Scott (1963) conclude that prolonged exposure to a salient object is crucial for the formation of attachments in mammals. Scott (1963) further theorized that emotional arousal in the presence of a salient object facilitates the precess of attachment formation. In a longitudinal study of the development of attachments in human infants, Schaffer and Emerson (1964) found that the strength of the attachment to the mother at 18 months of age was positively related to the extent of maternal-infant interaction and the responsiveness of the mother to the infant's cry. Since prolonged exposure and emotional arousal are conditions which typically characterize the human maternal-infant relationship, it is not surprising that Schaffer and Emerson found in all sixty of the children in the sample, evidence of a specific attachment to the mother by 18 months of age and in 90% by 11 months of age.

In a recent review of maternal-infant relationships in mammals, King (1966) theorized that the distress which would follow the exposure of a solitary infant to novel stimuli is inhibited in the presence of the mother because of the positive affective state which she evokes in the infant. The work of Harlow and Zimmerman (1959) with rhesus monkeys raised with cloth and/or wire surrogate mothers suggests that a particular type of prior stimulation arising from interaction with the mother object is required in order for her presence to have a distress inhibiting function. It is postulated that without the type or prior experience which results in a specific attachment, the mother's presence will not have a distress inhibiting effect. The importance of attachment related experience is supported by Rheingold's (in press) study which showed that the presence of the mother in a strange room inhibited distress in 10 month old humans whereas the presence of a female adult stranger resulted in a display of distress comparable to being alone. Also, a study of Morgan and Ricciuti (in press) indicated that the approach of a strange adult evoked distress in 12 month old humans who were seated several feet from their mothers; however, this distress was significantly attenuated when the infants were on their mothers' laps. Thus for human infants the proximity of strange adults, rather than inhibiting distress, may be the occasion for the evocation of distress.

Experimental evidence is abundant and consistently in support of maternal inhibition of distress in humans below three years of age (see also Arsenian, 1943, and Ainsworth and Wittig, in press), and clinical evidence (e.g. Freud and Burlingham, 1943) suggests that this effect continues into later childhood. However, there appears to be no experimental evidence for children over three years of age which bears directly upon the existence of a unique role for attached individuals as inhibitors of distress.

In this study children were exposed to a fear stimulus either in the presence of their mothers or in the presence of an adult female stranger. Ratings were made of their facial expression, movement about the room, and visual orientation over a five minute period. It was predicted that these behavioral indices would differ between the subjects in the two experimental conditions in a manner which would suggest less initial fear and a more rapid dissipation of fear for subjects in the presence of the mother.

Method

Subjects were obtained by writing letters to the parents of children enrolled in a private nursery school near Syracuse University. The letter described the experimental procedure in detail and was followed by a phone call, during which an individual appointment was made. Sixty per cent of the mothers contacted made and kept appointments. Reasonably complete data were obtained for 16 children who ranged in age from 3 years 7 months to 5 years 9 months with a mean age of 4 years 6 months. Their fathers had an average of 16.9 years of education and their mothers had an average of 14.7 years.

Fear stimulus. The fear stimulus was a toy mechanical gorilla 10 inches tall and covered except for the molded plastic face and hands with a gray-black plush fabrick (formerly distributed by Louis Marx Toy Co., Inc., New York City). The gorilla is operated from a remotely situated 3 volt DC power source and current reversing switch. Depending upon the direction of current flow the gorilla will either walk slowly in a straight path or stand still, turn his head from side to side, and beat his chest. These movements are accompanied by the clearly audible sound of the internal motor.

Prior to its exposure the Gorilla was concealed in a black box. The front side of the box was hinged at the base and held closed by an electro-magnetic catch. (Fig. 2) The wires from the gorilla and the magnetic catch passed out of the rear of the box and to an adjacent observation room.

The test room was 9 by 15 feet with a reflecting one-way vision glass on one of the 9 ft. walls. The box containing the gorilla was placed on the floor below the observation window. To the right of the gorilla box and angled across the corner of the room was a bench containing a toy cash register and a tiered plastic marble race toy. There were two chairs at the opposite end of the testing room, one in each of the corners and equi-distant from the gorilla box. All other items which seemed attractive to children had been removed from this functionally windowless room.

Procedure

The experimental procedure consisted of an adaption phase and a test phase. The adaption phase was the same for all subjects. When subjects arrived for the experiment they were met by the female assistant who told them that Dr. Schwarz was occupied but would be free to talk to the mother shortly. The assistant offered to wait with them in the test room, stating on the way that she knew some games to play while they waited. The assistant began the first game by tossing a bean-bag to the mother and encouraging her to toss it to the child. After two minutes of "bean-bag" the assistant initiated "follow-the-leader". After 4 minutes of play the mother took a seat on the right hand side of the room while the assistant administered a picture vocabulary test to the child. Next, at 5 minutes esapsed time, the assistant lead the child to the marble race and cash register which had been covered by a sheet. She demonstrated their operation and encouraged the child to play.

At this point the experimenter entered and introduced himself to the mother. A flip of a coin had determined the assignment of experimental conditions. If the child were to be in the Mother Condition, the Experimenter asked the assistant to come out of the room, and remarked to the mother that he would see her shortly. For the Assistant Condition the experimenter asked to see the mother and, before leaving, turned to the child and remarked casually that his mother would be back shortly. The assistant took up the mother's position. Both mother and assistant were given a magazine containing instructions to ignore the child and the gorilla, to respond to the child's question with a brief, "I don't know", and if the child persisted to remark, "I have to read this magazine".

The departure of the mother or the assistant marked the beginning of the test phase. Observers behind the one-way glass began recording aspects of the child's behavior and continued for the next 6 minutes. While playing with the toys the child was about 2 ft. from the gorilla box. After 60 seconds of play a switch was thrown in the observation room to start the gorilla beating his chest and turning his head from side to side while in the box. Then the magnetic catch was released causing the front of the box to flop down with a slap which invariably attracted the child's attention. Now the gorilla was exposed to the child's full view. The chest beating was continued for a few more seconds and then the current was reversed, causing the gorilla to walk slowly and ponderously down the ramp. When the gorilla had reached a position about 17 inches from the box, the current was reversed, again causing the gorilla to stop walking and to resume beating its chest and turning its head. After a total of about 25 seconds all action of the gorilla was stopped and it remained stationary for the remainder of the test phase. Observations of the child with the conditions unchanged continued for an additional 4½ minutes.

At the conclusion of the test phase the child was given a chance to operate the gorilla. Any misconceptions of the child were corrected by explanation and demonstration. Each child was brought to the point where he would willingly touch and hold the gorilla before leaving the experiment.

Behavioral observations and ratings

Facial expression. As one method of assessing the degree of apprehension aroused by the gorilla, ratings were made of the child's facial expression. A composite picture was developed which provided photographic examples of children's facial expressions representing seven points on a continuum of pleasure ranging from very unhappy or displeased to very happy. Three or four pictures which had been judged relevant to this dimension and for which the judges' ratings were low in variability exemplified each point on the scale. Every 15 seconds a light flashed in the observation room signalling the moment for a rating. The rater noted the child's expression of the moment, glanced at the pictorial composite, and assigned a number associated with the facial expression most similar to that of the subject at that moment. If the child was facing away from the rater, the rating was skipped. In the analysis of the data ratings were averaged for each minute of the test phase.

Unpleasant facial expressions were assumed to be indicative of fear and apprehension in this situation. Therefore, it was predicted that the facial expressions of S̲s in the Assistant Condition would indicate greater displeasure at the exposure of the gorilla and would return to the neutral or pleasurable portion of the scale more slowly than the facial expressions of S̲s in the Mother Condition.

Position in the room. The floor surface of the testing room was marked off into 18 inch squares to facilitate recording of the S̲s position in the room. With the flashing of the signal light every 15 seconds in the observation room an observer noted the S's position in the room. From this record it was noted whether the S left the area of the toys and the gorilla and the time at which he returned.

Since the person in the room ignored the S, and since the toys near the gorilla were the only attractive item in the room, it was assumed that the S's distance from the area of the toys and the gorilla would reflect his degree of apprehension or fear of the gorilla. Accordingly, it was predicted that Ss in the Mother Condition would be less likely to leave or would return more rapidly to the area of the toys and the gorilla than Ss in the Assistant Condition.

Visual Orientation. A multiple event recorder was used to record the child's visual orientation. Four categories of response were recorded: looking at the gorilla, at the toys, at the mother or assistant (person), and looking elsewhere about the room. This equipment was not available at the start of the study, therefore records were obtained for only 10 subjects, 5 in each condition. No specific predictions were made for visual orientation.

Results

The data on facial expression for a given S consisted of six scores: the average rating for the first minute of the test phase before the gorilla emerged, and an average rating for each of the five succeeding minutes of exposure to the gorilla. These data were analyzed in a mixed Treatments by Subjects (Lindquist Type I) analysis of variance design. The interaction effect between treatment conditions and trials was statistically significant. ($F = 3.14$, $p < .02$, $df = 5/70$). The trend over trials was not significant, for the assistant Condition, but was significant beyond the .001 level in the Mother Condition ($F = 7.94$, $df=5/35$).

Contrary to prior prediction, children whose mothers were present had a sharp decline in the pleasantness of facial expression as a consequence of exposure to the gorilla, where as the children in the presence of the female assistant showed no significant change in expression. The difference between the means of the two groups was greatest during the first minute of exposure ($t = 2.38$, $p < .05$) and decreased progressively as a result of increases in pleasantness of expression among subjects in the Mother Condition.

The other index of apprehension, time to return to the area of the toys and the gorilla, also yielded differences between the two conditions which were opposite to prior prediction. It took longer for children in the Mother Condition to return to the area of the toys and the gorilla; however, this difference only reached the .10 level of significance with the Kolmogorov-Smirnov two-sample test ($N = 16$).

Significant biserial correlations were found between ratings of facial expression during exposure to the gorilla and whether the child returned or not to the area of the gorilla and the toys by the end of the five minute exposure period (See Table 1). Unpleasant facial expressions during gorilla exposure were associated with failure to return to the toys. Expression prior to gorilla exposure was not significantly related to the child's return to the toys.

Since the number of Ss for whom the visual orientation data was available is small, conclusions which may be drawn from this data are limited. Table 2 presents the mean percentage of total time spent looking at four categories of objects during the 5 minutes of exposure to the gorilla by children in the Mother and Assistant Conditions. The children with their mothers spent more time looking at the gorilla than those with the assistant, however, the Mann-Whitney test of this difference only reached the .08 level of significance. This surplus time spent looking at the gorilla by the mother group seems to have been taken away from the categories of Toys and Person.

In Figure 2 the mean number of seconds spent looking at the four objects during each minute of the test phase for Ss in the Mother and Assistant Conditions are plotted as a function of time. Analysis of variance (Lindquist Type I) revealed significant main effect for Time for all objects except Person. The experimental treatment effect was not significant for any object nor was the interaction; however, the interaction effect reached the .10 level for Gorilla.

A correlational analysis indicated, in general, that Ss distant from the gorilla had unpleasant facial expressions and spent more time looking around the room and less time looking at the toys than Ss close to the gorilla. But position in the room was not related to the time spent looking at the gorilla. Time spent looking at the gorilla was negatively related to the time spent looking at the toys.

Discussion

It is apparent that most of the results were opposite to what one would expect if the mother had had a greater distress inhibiting effect upon the child than the stranger. Yet it does not seem possible that the mother could have had a fear inducing effect upon the child. Her presence, however, may have facilitated the communication and expression of fear. Being left in a strange room with a strange person may have induced a general inhibition of motility and emotional communication. The assistant may have been an aversive stimulus. Her presence at one end of the room and the appearance of the gorilla at the other, perhaps generated an avoidance-avoidance conflict which was resolved by focusing attention on the toys, the only positively valued stimulus in the room. It may even be conjectured that the presence of the mother gave children in the Mother Condition sufficient security to cope with the frightening gorilla as manifested by their tendency to spend more time looking at the gorilla.

An alternative explanation involves the assumption that the presence of the assistant continued to evoke in the children a set or expectancy for play and amusement. With this type of psychological set the appearance of the gorilla, an ambiguous as well as a novel stimulus, may have been given a nonthreatening construction which was compatible with that expectation; namely, that the sudden appearance of the gorilla was another game perpetrated by the playful assistant. Some facts are not entirely consistent with this explanation, however. First of all, one would expect Ss in the Assistant Condition to have happier facial expressions in the minute before the appearance of the gorilla, which they did not. Secondly, the Ss in the Assistant Condition, for the most part, did not make playful responses toward the gorilla, but rather seemed to actively ignore it, after initial minute of intense visual inspection.

Since this study is inconclusive with regard to the existence of a greater distress inhibiting function for the mother as compared with a stranger in children from three to five years of age, other studies with additional experimental and control groups are planned. Since no children were exposed to the gorilla while alone, it cannot be concluded that the mother had no distress inhibiting effect. It is clear, however, that the effects of the presence of an attached individual as compared with an unattached individual are not as marked for children from 3½ to 5½ years as they are for children from 9 months to 2½ years.

References

Ainsworth, Mary D.S. and Wittig, Barbara A. Attachment and exploratory behavior of one-year-olds in a strange situation. In B.M. Foss (Ed.), *Determinants of Infant Behavior*: *IV*. New York: Wiley, in press.

Arsenian, Jean M. Young children in an insecure situation. *J. abn. soc. Psychol.*, 1943, 38, 225-249.

Cairns, R.B. Attachment behavior of mammals. *Psychol. Rev.*, 1966, 73, 409-426.

Freud, A. and Burlingham, D. *War and Children*. New York: Medical War Books, 1943.

Harlow, H.F. and Zimmermann, R.R. Affectional responses in the infant monkey. *Science*, 1959, 130, 421-432.

King, D.L. A theory of infant-mother relationship in mammals and birds. *Psychol. Bull.*, 1966, 65, 143-155.

Morgan, G.A. and Ricciuti, H.N. Infants' responses to strangers during the first year. In B.M. Foss (Ed.), *Determinants of Infant Behavior*: *IV*. New York: Wiley, in press.

Rheingold, Harriet L. The effect of a strange environment on the behavior of infants. In B.M. Foss (Ed.), *Determinants of Infant Behavior*: *IV*. New York: Wiley, in press.

Schaffer, H.R. and Emerson, Peggy E. The development of social attachment in infancy. *Monogr. Soc. Res. Child Devel.*, 1964, 29, No. 3, (Whole No. 94).

Scott, J.P. The process of primary socialization in canine and human infants. *Monogr. Soc. Res. Child Devel.*, 1963, 28, No. 1, (Whole No. 85).

Footnotes

*Psychology Department 331 Huntington Hall

Syracuse University

Syracuse, New York 13210

[1] This study was supported by Contract OE04120 United States Office of Economic Opportunity. The author wishes to express his appreciation to Linda L Dafidoff, Irwin L. Keppler, Sandra R. Diener, Cynthia H. Deitch, and Toni G. Hollander for their assistance in the collection analysis of the data. A special notice is due Mrs. Mildred Blackburn, Director of the Erwin Methodist Nursery School for her assistance in locating children for this study.

Figure Captions

Fig. 1. A plot of the means of the Mother and Assistant Conditions for facial expression ratings averaged over one minute periods from one minute prior to gorilla exposure through the fifth post-exposure minute.

Fig. 2. Graphs of the mean number of seconds spent looking at Person, Room, Gorilla, and Toys during each minute of the test phase for _S_s in the Mother and Assistant Conditions.

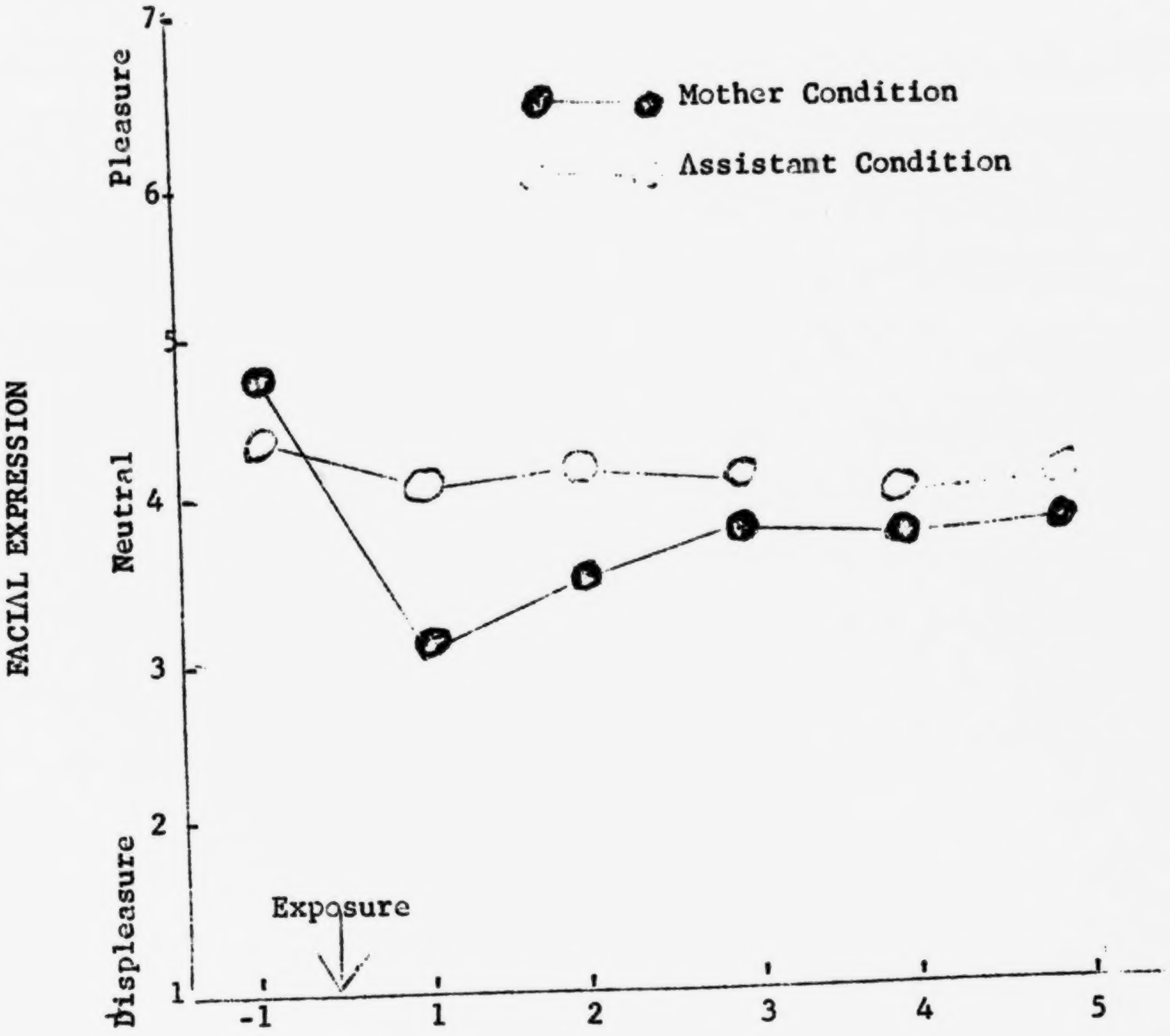


Fig.1. A plot of the means of the Mother and Assistant Conditions for facial expression ratings averaged over one minute periods from one minute prior to gorilla exposure through the fifth post-exposure minute.

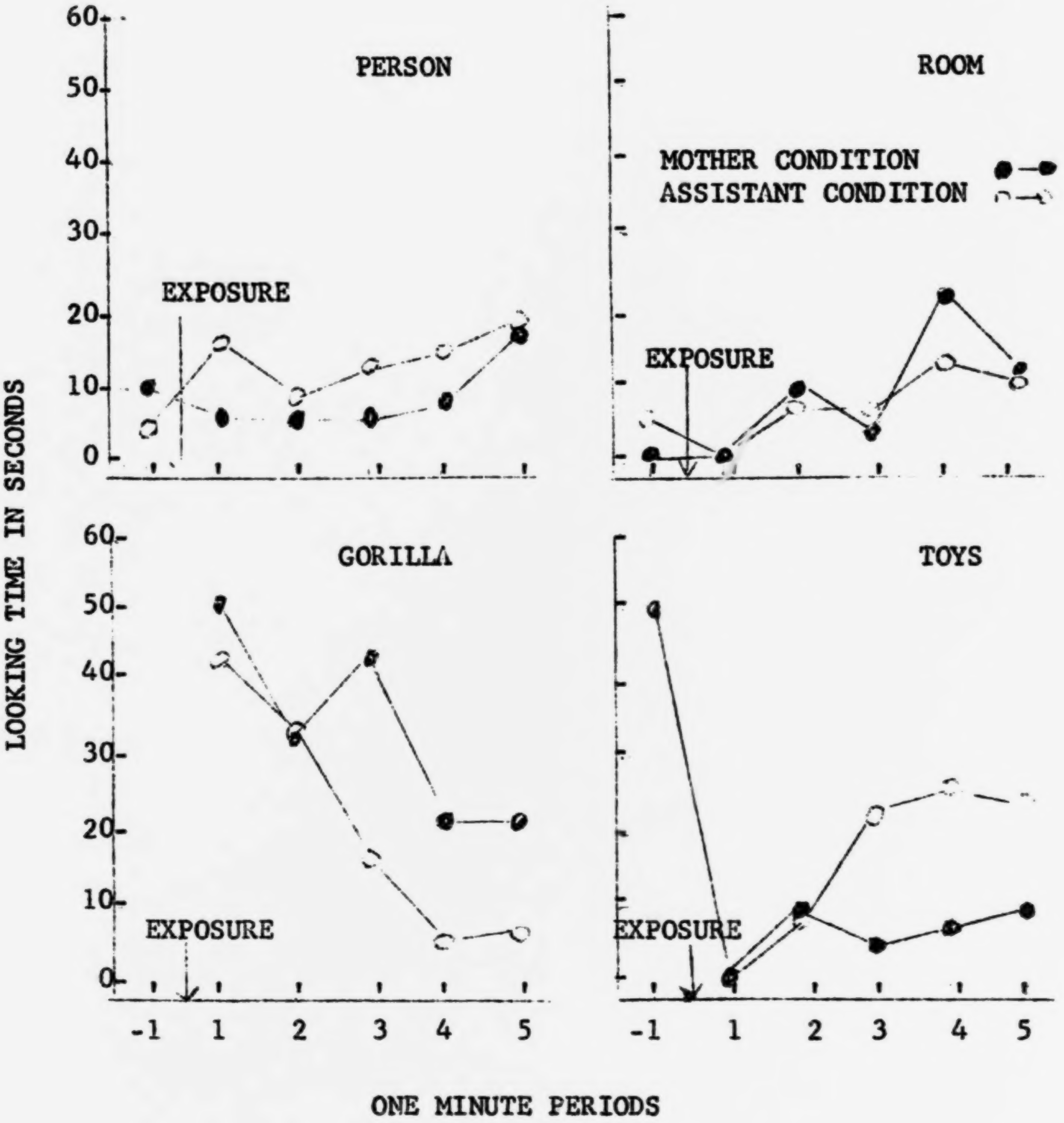

PERSON
ROOM
GORILLA
TOYS
MOTHER CONDITION ●—●
ASSISTANT CONDITION ○—○
EXPOSURE
EXPOSURE
EXPOSURE
EXPOSURE
LOOKING TIME IN SECONDS
60
50
40
30
20
10
0
60
50
40
30
20
10
0
-1
1
2
3
4
5
-1
1
2
3
4
5
ONE MINUTE PERIODS

Table 1

Biserial correlations between Average Facial Expression for each minute of the Test Phase and the dichotomous classification of Return (N = 9) vs. Nonreturn (N = 7)

One Minute Periods					
-1	1	2	3	4	5
-.40	.73*	.83**	.64*	.58	.57

* p = .05 two-tailed test

** p = .01 two-tailed test

Table 2

Mean percentage of total time spent looking at four categories of objects during the 5 min. postexposure period by children in the Mother and Assistant Conditions

	GORILLA	TOYS	ADULT	OTHER
Mother Condition (N = 5)	57%	13%	14%	16%
Assistant Condition (N = 5)	36%	27%	24%	13%
p-value of the Mann-Whitney U	.08	.34	.11	.21

Evaluating Behavioral Change During a Six-Week Pre-kindergarten Intervention Experience

Lois Hayweiser, David Massari, and W.J. Meyer[1]

This project grew out of a unique set of circumstances occurring in the City of Syracuse and Onondaga County during the summer of 1967. Essentially, there occurred a number of disagreements between the local Community Action Agency and the Office of Economic Opportunity making it unclear whether or not a Summer Head Start program would be funded. In addition it seemed apparent that the community of professional child development specialists would have minimal involvment in the Summer Head Start program. One consequence of the discord was the refusal of several communities in Onondaga County to submit proposals for Summer Head Start programs. Nevertheless, at least one of these communities was sufficiently concerned about their "culturally deprived" children that we were contacted and asked if we could help support a six-week pre-kindergarten program for approximately forty-five children. After receiving approval from the Institute for Educational Development (IED) to expend funds for such a project, and after agreement by the school officials that we could perform an extensive evaluation of the program and use the children for other research purposes, we agreed to undertake the project.

Responsibility for the content of the six-week program was given to the teachers and aides. Our primary objective was the development of evaluation techniques and the assessment of these techniques against certain standard procedures, some of which were used in the national evaluation of Project Head Start.

---

[1] The authors wish to express their appreciation to the entire staff of the Syracuse Evaluation and Research Center for their contributions to this project. A particular note of appreciation is due to Mrs. Jacquline Morrow Massari who contributed many ideas to this project and who served, on a voluntary basis, in the collection of the data.

Thus, instruments such as the Zigler Behavior Inventory and the Stanford Binet were included as part of the evaluation program. Of equal concern were the changes in the children's behavior during the six-week period. Since the children continue their education in the same School District, it also will be possible to maintain a longitudinal assessment of their accomplishments.

Because of the extensive number of measures collected for each child and the consequent enormity of the data analyses, it was impossible to complete the data analysis in time for this report. The salient features of our findings are included and focus on those issues of primary concern to us. Subsequent reports will provide additional analysis.

Several considerations provided direction for the instruments employed in the evaluation. There was every reason to expect that the six-week experience would bring about significant gains in Stanford Binet performance. The reasons for the gains, and their meaning in terms of changes in intellectual competence, were unclear. Thus, it was our hypothesis that the higher I.Q.'s typically found on post-test Binet performance do not reflect change in cognitive structure. Our position, which is consistent with that proposed by Zigler (1967), is that observed increases in I.Q. reflect change in some aspects of the child's motivation. More specifically, it was assumed that children are more willing to emit responses, both verbally and motionally, because of an increased confidence in adults derived from positive encounters with adults during the intervention experience. One aspect of our assessment procedure, therefore, was to evaluate changes in the children's willingness to emit responses to the various cognitive demands of the Stanford-Binet.

Another aspect of our assessment of cognitive behavior focused on "cognitive style". Studies by Kagan, Moss, and Seigel (1963), Maccoby, Dawley, Hagan, and Deguman (1965) and Hess and Shipman (1964) clearly suggest that children approach cognitive tasks differently. Maccoby et al. (1965) examined cognitive style in terms of a lack of "motor inhibition" and reports a positive correlation (r = .41) between her task, Draw A Line Slowly, and Stanford-Binet I.Q. It would appear that a lack of "motor inhibition", or "impulsivity", is a general style pervading cognitive behavior and which may also influence social behavior.

In descriptive terms, the impulsive child seems to respond to stimulus situations at a too rapid rate in order to process the information presented . Components of the stimuli may be missed altogether or responses may be made only to the most dominant or familiar features of the stimuli. There is the additional possibility that the child does not even consider the stimulus characteristics of the problem but responds on some totally irrelevant basis unique to the child. Impulsive children, whether the antecedents are biological, environmental, or a combination of both, find complex tasks, which require processing time, more difficult. Thus, lower performance on a measure of intelligence such as the Binet would be anticipated.

The inability to inhibit responses should also negatively influence how teachers view the adequacy of the social behaviors of impulsive children. The failure to "hear" directions, excessive wiggling at the desk, or running in the hallway, are behaviors generally viewed as negative by teachers. It is just such behaviors, however, which apparently characterize the "impulsive" child.

A second broadly defined area of the evaluation focused on the adequacy of the children's social behavior. Of concern here is the relationship between the teacher's perception of the adequacy of the child's social behavior and Binet gain.

Our assumption was that children who conform more to the demands of the classroom situation are preferred by teachers and they therefore provide them with a more conducive learning environment. Rosenthal (1967),for example, shows that the teacher's initial perception of her pupils influences the amount of gain reported. A scale concerned with children's adaptive classroom behaviors was developed, in collaboration with the teachers, and administered on a pre- and post-test basis. At no time were the teachers, or the aides, aware of the child's Binet I.Q. so that all ratings can be considered independent of any obvious biasing effect.

A third consideration in designing this project was an examination of the teachers' as opposed to the aides' perceptions of the children. One might speculate that teachers view children differently from aides because of their more academic orientation. Aides, it might be presumed, accept this role for a variety of motives, one of which is most likely to be that they enjoy children. One might anticipate, therefore, that the aides' perception of children is less influenced by their estimate of the child's intellectual ability, or by his level of achievement. An alternative reason for expecting differences derives from the fact that aides generally are not as experienced in judging groups of children in a classroom situation as are the teachers. Since the three aides for this project were sophomore or junior students at state teacher colleges in upstate New York it would be expected that any differences would more likely be based not so much upon their differences in academic orientation as their lack of an experiential standard for a classroom type situation against which to compare the children.

A final consideration in the design of this project was to provide an assessment of the effects of a six-week intervention experience in comparison with children from the same population not having the experience. In this connection, more children were invited to participate in the program than we anticipated would finally attend. Unfortunately, in some respects, this procedure worked better than we had hoped for in that more children than anticipated either failed to attend altogether or attended for only one or two days. Fortunately, it was possible to assess the general level of intellectual functioning of these children who had been invited to attend the program but who, for a variety of reasons, did not do so.

In summary it can be stated that our major focus was on social cognitive components related to behavioral change. Although our interest was focused somewhat on demonstrating change as a function of the intervention experience, our primary concern was with the relationship of cognitive, social, and teacher-aide variables and observed behavioral change.

Method

Subjects: The subject pool from which the sample was drawn was defined by the Director of Elementary Education (Mr. Paul Anderson) of the Liverpool School District. This school district includes a broad range of socio-economic levels. The sample identified met the poverty criterion with respect to family income and were further known to the school district officials either through social service agencies or through prior encounters with the family because of problems arising with older children in the family. Of the approximately 85 children identified, 45 children were selected as being in greatest need for a pre-kindergarten program. The parents of all 45 of these children were personally contacted by the school officials and the program described to them. A total of 33 of the invited children finally attended the pre-school. Characteristics of these children are summarized in Table 1 from which it can be noted that their pre-test Binet scores are not

Table 1

Characteristics of Sample

	CA(Mos.)	Binet	Zigler
M	64.8	90.0	138.5
SD	6.9	16.8	20.6

unlike those typically reported in other Head Start classes. The somewhat higher range of chronological ages reflects the inclusion of two children who had been denied entrance to the first grade after a kindergarten program on the grounds that they were insufficiently ready for that experience.

Staff

The three teachers were comprised of two females and one male. Prior teaching experience was 3, 7, and 10 years. Each of the three aides, all females, were enrolled in college and were in teacher preparation programs.

All Stanford-Binets were administered by Examiners with considerable experience in administering the Binet to young children. Other cognitive assessments and classroom observations were conducted by a core of 8 advanced level graduate students.

Instruments

A. Cognitive Measures.

The Stanford-Binet, Form LM, was administered on a pre- and post-test basis. This test, rather than some other measure, was selected simply because it has been used in other similar studies making comparisons with our data easier. Each of the Examiners gained extensive experience with the instrument and the age level because of their having examined the children in our Head Start Evaluation sample. Children were randomly assigned to Examiners. In order to maximize reliability within measurements, the same Examiner was used for each child on the pre- and post-tests. Independence between measures was assured by using different Examiners for the different measures. Standard manual procedures were followed.

The second measure focused on the children's responses to the cognitive demands imposed by the Stanford-Binet. More specifically, children's responses were rated by the Examiner during the testing, in terms of "work" or "non work" following a procedure adapted from Birch (1967). The scales are: Work-No Work Categories (Adapted from Hertzig, Birch, Thomas, and Mendez, 1967).

A. Work Responses (child must attempt what was asked of him)

1. Work Verbal Spontaneous Extension -- S spontaneously extends his response. For example, in completing the analogy "birds fly and fish ____" the child says "bite-like one once bit me on the leg."

2. Work Verbal Delimited -- S's response is restricted to defined requirements of task.

3. Work Non-Verbal Spontaneous Extension -- S responds to demands of task and then spontaneously performs some additional behavior. For example, child strings beads and then places around her neck.

4. Work Non-Verbal Delimited -- S responds strictly to the demands of the task.

B. Non-Work Responses (Child does not work at task)

1. Non-Work Verbal Competence -- S refuses to attempt task stating a lack of competence to handle task. For example, "I am too young for that."

2. Non-Work Verbal Negation -- S verbally refuses to work. For example, "I won't."

3. Non-Work Verbal Substitution -- S offers an irrelevant verbalization such as "I want to leave now," or "I want to play with the toys."

4. Non-Work Verbal Aid -- S verbally requests aid from examiner.

5. Non-Work Non-Verbal Negation -- S refuses to work by indicating with a motor response. For example, head shake, pushing materials away, or turning away from Examiner.

6. Non-Work Non-Verbal Substitution -- S engages in an irrelevant motor response. For example, when asked to build a bridge, the child walks away

or pushes the blocks back and forth on table.

7. Non-Work Non-Verbal Passive -- S does not respond in any way to task or Examiner.

Prior to testing the sample children, each Examiner administered a minimum of three Binets to preschool age children with two other raters present. Observer reliability was estimated in terms of the per cent of instances where all three observers were in agreement with respect to the 11 categories. The median agreement was approximately 74%. Agreement in terms of the broad categories of "work" "non-work" was approximately 90%. On the basis of these data, it was felt that the Examiners could rate the "work" "non-work" responses of the children while administering the Binet test.

The "work" "non-work" ratings were included to measure change in children's responses to cognitive demands under, at least, somewhat stressful conditions. Partially, these measures should reflect the child's willingness to attempt responses where the probability of failure is high. Assuming that the child responds correctly to some items that were, in various ways, avoided on the pre-test, it may be possible to account for gains in Binet I.Q. The "work" "non-work" measure may also reflect a child's "style" in approaching difficult tasks. Thus rather than attack a difficult task, a child may prefer either not to respond at all or respond in a completely irrelevant fashion. Finally, the "work" "non-work" measure may also reflect a motivation to, in a sense, please the Examiner.

Several impulsivity measures were used in this study. The "Draw a Line Slowly" (DAL) Test, reported by Maccoby, et al. (1965) and Hess, et al. (1966) was used with certain modifications. This task requires the child to draw a line, beginning at the top of a plain 8½ x 11" piece of paper and proceeding to the bottom of the page, as slowly as possible.

The following instructions were used: "I am going to draw a line on this paper as fast as I can. I will start here at the top and go to the bottom. Now you try it, take the pencil and go as fast as you can to the bottom. Very good. Now, this time I am going to draw a line from the top to the bottom of the page as slow as I can. Watch. Now you try it. Draw the line as slow as you can. That was very good." These were practice trials and to show and explain to the children the meanings of "slow" and "fast". The children were then given the test trials in the following sequence of conditions: 1. "Draw the line slow"using an 8½" x 11" piece of paper with no markings on it. Here the direction was to draw the line as slow as possible starting at the top and going to the bottom. 2. "Draw the line slow" using an 8½" x 11" piece of paper on which there was an X at the top and bottom. The child was told to draw the line "even more"slowly and to connect the Xs. 3. This condition repeated number two. 4. "Draw the line fast" using an 8½" x 11" piece of paper with no markings. 5. "Draw the line fast" and connect the X at the top to the X at the bottom.

The second impulsivity measure was the Walk a Line Slowly Test (WAL). In this task two six feet long parallel lines of adhesive tape, five inches apart, were placed on the floor. The child was instructed to place one foot on each tape. The task was given under three conditions, in the following order: 1. "I want you to walk to the end of the tape making sure you do not step off the lines." No instructions concerning speed were given. 2. "Now I want you to walk the line as slow as you can." 3. "Now I want you to walk the line as fast as you can."

Two measures, not previously reported by other investigators, were used. One measure was an adaptation of the "Perceptual Speed Test"(PST) subtest from the Primary Mental Abilities Test, Primary I Battery. This test requires the child to match a standard against four alternatives. In addition to using the standard scoring system of recording number of correct responses, a latency to first response measure was used and a correction procedure in which the children continued responding to each item until making a correct choice. A short latency was assumed to indicate a lack of careful analysis of the response alternatives (One problem with this assumption is that a very bright child might be able to rapidly process the alternatives giving the impression of impulsive behavior.) In terms of responses-to-correct-choice measure, it was assumed that a child requiring many responses over the test items is responding more or less at random without considering the pertinent components of the stimuli.

The PST was administered following the verbal directions given in the manual. For the correction procedure, the children were told: "No, that is not the correct one, try again." "That one is wrong too, try again," and so forth. The more difficult items were (those appearing on page 13 of the test booklet) in order to reduce familiarity with the stimuli.

The final cognitive measure involved a modification of two WPPSI mazes, numbers 2A and 3A. Each maze was administered as it appears on the test and then with the blind alleys deleted. Following a practice maze, (Maze 1A from the WPPSI) each subject was given the test mazes in the following sequence: 1. no cul-de-sac, 2. easy maze with cul-de-sac, 3. no cul-de-sac, but longer than maze 1, and 4. harder maze with cul-de-sac. This task was designed to measure impulsivity with a more difficult paper-and-pencil task than was involved with the DAL. Performance was assessed in terms of errors, defined as crossing lines and/or entering a cul-de-sac (mazes 2 and 4, only), and elapsed time to complete the maze.

Social Behavior Measures

The main objective of the instruments described in this section is to provide data relevant to the adequacy of the children's social behavior as perceived by the teachers and aides. Although there was concern for behavioral change, our analysis also focused on the correlates of perceived social adequacy.

Since teacher perceptions probably effect, perhaps to a significant degree, the school environment of the child, it seemed reasonable to ask teachers to describe the behaviors influencing these perceptions. Each of our three teachers were asked, in separate interviews, to describe a maximally adapted and maximally maladapted kindergarten child. All statements were probed until descriptions were given in reasonably objective behavioral terms. For example, the statement "is well behaved" after probing resulted in statements such as "does not grab", "waits in line and stays in his place", or "waits for directions before rushing in". A total of 62 such statements comprise the Adaptive Behavior Rating Scale (ABRS), (see Appendix A for a copy of the scale.)

The teachers and aides completed the scale for each of 33 children. Both pre- and post-test measures were obtained. An index of internal consistency was obtained using a procedure described by Flanagan (1965) resulting in a reliability estimate of .86.

A second measure of teacher perception involved a man-to-man rating technique adapted from procedures described by Gardner and Thompson (1956). Essentially, this technique involves naming a child, from some population which includes the group to be rated but which is larger than that group, who has the most and least of the attribute to be subsequently rated. A third point is identified in terms of the most average child. All members of the group are then rated in terms of this reference population. Separate reference scales are established for each rated attribute. Each rater defines the children for his scale providing, at least theoretically, scales across raters that have psychological equivalence.

Two hypothetical situations were constructed and rated by both teachers and aides. Situation 1 was: "Suppose you were asked to select children for a special kindergarten class where academic superiority was important." Situation 2 was: "Suppose you were asked to select children for a kindergarten class in terms of their qualities as children that teachers enjoy." Ratings were made on a five point scale.

A third instrument used to assess social-emotional behavior was the Zigler Behavior Inventory. This instrument was employed in the national evaluation of Project Head Start and its properties are generally well known.

Finally, our assessment of social behavior included a very limited number of observations on approximately one-half of the sample. These children were randomly selected prior to the first day of classes and were observed for ten randomly selected one-minute periods over two mornings. This procedure was followed during the first and last sixth weeks of the program.

Categories of behavior were broadly defined and include the following:

1. Conformity-non-conformity: Observers judgment that child was doing what was expected, or not, during observation time;

2. Non-verbal communications: Child points, shoves, signals;

3. Incomplete verbal communications: Child uses one word or fragmented sentences;

4. Complete verbal communications: Child uses complete sentences or phrases;

5. Attentional changes: Child shifts attention from one task to another,

shifts from one situation to another (person entering room).

Observer agreement was determined by having the three observers observe and simultaneously code the behavior of similar aged children in one of the city preschools. These practice sessions were done during both structured and free play situations for a number of children. Practice was continued until the raters achieved 90-100% agreement for the number of responses in the categories of conformity-nonconformity, verbal vs. non-verbal and 70-80% agreement for attentional changes and incomplete vs. complete verbal communications.

Classroom and Home Stimulation

Two measures of input were secured on each child in the sample. In an effort to assess the level of stimulation of the child's home, the teachers accompanied by an aide or research assistant, made home visits using the Inventory of Home Stimulation (STIM) developed by Caldwell (1967). This instrument assesses such variables as orderliness, number and kinds of play materials and verbal interactions of the mother. A copy of our modification can be found in Appendix A.

The classroom environment was assessed in terms of the teachers and aides use of "praise" and "blame". We were especially interested in assessing the initial impact of the teachers on the children so that 30 minutes of observations were made for each teacher and aide over the first hour of the first day of classes. Teachers and aides were observed in alternate periods of 10 minutes each. Praise, blame, object and situation content were recorded. An additional one-half hour was obtained for both teachers and aides on one additional day during the first week and on one day during the last week of the program. Further observations were made of both teacher and aide during a field trip to a farm and during one-half hour of the lunch period.

Observer agreement was obtained by having the observers simultaneously rate two teachers in a city nursery school for two five minute periods. Observer agreement was between 90 and 100% for praise or blame, the object and essential content of the situation.

Results

The results are organized in three general categories: 1. behavioral changes during the six-week program with three subsections consisting of cognitive changes, changes in social behaviors, and changes in teacher behaviors; 2. the interrelationships among the variables; and 3. a summary of two special projects within the overall program concerned with arithmetic skills.

Although the various indices reflect specific theoretical biases, it should be made clear that this is an exploratory study. The Syracuse Center is currently attempting to cross-validate some of the more promising relationships and hope that other investigators may wish to work with some of the measures. Until such validating studies are conducted the measures should not be used for other than research purposes.

One additional comment may serve to reduce reader confusion. Obviously our measures were obtained over several days and were scheduled with phenomenal precision. An absent child often could not be rescheduled so that the N varies somewhat over analysis. Most of the change score and correlational analyses are based on N's of between 29 and 33 children. All instances of an N below 29 will be specifically noted.

A. Effects of Intervention Program

1. Cognitive Behaviors.

All analyses in this section, unless otherwise noted, consist of correlated _t_ tests following standard procedures. Summarized in Table 2 are the means, SDs,

Table 2
Summary of Pre-Post Means, SDs and t tests

	Stanford-Binet		PST (Responses)		PST (Latency)		DAL Rates	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
M	90.0	96.1	18.3	15.7	44.3	44.8	2.0	2.2
SD	16.8	19.7	4.0	4.1	12.5	22.2	1.9	1.9
	t = 6.01**		t = 3.96**		t = <1		t = <1	

	Mazes (Time Scores)		Mazes Errors		WAL		Teacher Academic Ratings	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
M	60.4	50.7	7.2	5.4	10.5	11.0	1.1	2.6
SD	17.1	12.7	4.5	4.4	6.7	4.6	.8	1.2
	t = 2.5*		t = 1.8		t = <1		t = <1	
							df = 30	

	Aides Academic Ratings	
	Pre	Post
M	2.7	2.6
SD	1.2	1.4
	t = <1	

** $p < .01$
* $p < .05$

and *ts* for each cognitive measure. The average gain in Binet I.Q. was 6.1 and is statistically significant. This average amount of gain is certainly not spectacular but is not inconsistent with gains reported in other intervention studies. Consistency and relative position between pre- and post-test is remarkably high as reflected in a product moment correlation of .98. As anticipated, there is a significant relationship ($r = .52$) between pre-test IQ and change score showing that the brighter children made the greatest IQ gains.

Performance on the Perceptual Speed Test (PST) also improved significantly over the six-week period. It should be noted that improvement on this measure is denoted by a reduction in the average score which reflects, because of the correction procedure used, a reduction in the number of responses made to each item before selecting the correct response. This result suggests that the children were somewhat more careful in making their choices on the post-test, but it cannot be argued that they were necessarily taking more time in their responses. This latter argument is supported by the very slight gain in latency to score in making their initial choice on this task. Relative position with respect to number of responses made on the pre- and post-tests was a moderately high relationship as reflected by a product moment correlation of .68. The relationship between pre-test performance and change score is lower than that observed for the Stanford-Binet ($r = .39$), but statistically significant. The lower correlation probably reflects a ceiling effect in that children obtaining an initially low score could only improve by one or two points. Again, the correlation suggests that the intervention experience is more profitable for those children with initially high scores. The test-retest correlation for the latency measure on the PST indicates essentially no relationship between the two measures ($r = -.14$, $df = 27$, $p = >.05$). Undoubtedly this correlation reflects the

Table 3
Work Non-Work Behaviors

Total Performance

Work Verbal Items		Work Performance Items	
Pre	Post	Pre	Post
83.1%	85.0%	92.8%	97.0%
t = .90 using arc sin trans		*t* = 2.8 using arc sine	

Ceiling Year Items

Work Verbal Items		Work Performance Items	
Pre	Post	Pre	Post
69%	74%	75%	96%
T = 117 (Wilcoxon Matched Pairs)		T = 0	

general problem of obtaining reliable latency measures, but as will be seen in the next section of this report the post-test latency measure, but not the pre-test latency measure, is related to performance on the Stanford-Binet.

Analyses of the Draw-A-Line Slowly (DAL) Test is based on a rate measure; that is, length of line divided by time to draw the line. (Analyses involving the straight latency measure yield identical results.) Thus in this measure a high score is indicative of impulsivity. Scores on the three trials were highly intercorrelated (median r = .89) therefore scores for each child were pooled. It will be noted from Table 2 that the change in performance from pre- to post-test was not statistically significant. There was, however, considerable consistency in relative position on the pre- and post-test as reflected in a product moment correlation of .84. The correlation between pre-test performance and change score was .29 again indicating a tendency, which is not statistically significant, for greater improvement to occur among the initially better performers on this measure. That this correlation is low may reflect a ceiling effect in that initially low performance by some of the children could just not be improved upon. It is noteworthy, however, that among the high impulsive children there was apparently little general gain. A similar result was found on the Walk-A-Line Slowly Test where again the difference between the pre- and post-test measures was not statistically significant. The measure of consistency for the WAL was lower (but statistically significant) than that found for the DAL (r = .41). There is no apparent reason for this difference unless, in fact, the WAL is a less reliable measure. The correlation between pre-test performance on the WAL and difference score is negatively related (r = -.55) indicating that initially slow performance is associated with slower (less impulsive) performance on the post-test.

Performance on the WPPSI mazes in terms of pre- and post-performance indicated a significant increase in the speed with which the tasks were performed but there was no significant decrease in the number of errors (the decline in the number of errors approached statistical significance, $\underline{t} = 1.8$; $\underline{df} = 18$; $p = < .08$). These results suggest that when children are not given specific instructions to perform a task slowly, there is a tendency, at least with respect to the maze task, to speed up performance. The correlation between error scores on the pre- and post-tests was $r = .50$ and the relationship between pre-test score and change score was $r = .49$. With respect to the latency measure for the mazes, the correlation between pre- and post-test performance was .33, which for 19 $\underline{df}$ is not statistically significant, and the correlation between pre-test performance and change score was $r = -.73$.

One of the primary concerns in this project is the teachers perception of the children. Recall that we administered an adaptation of the Syracuse Scales of Social Relations asking each teacher and aide to rate each child with respect to academic capability. Examination of the results of this procedure, shown in Table 2, indicate that neither the teachers nor the aides felt that there was significant change in the youngsters over the six-week period. It should be noted that, on the average, both the teachers and the aides rated the children to a considerable degree at the low end of the scale. As anticipated, there was a significant tendency for the aides to rate the children higher on the pre-test than did the teachers ($\underline{t} = 7.8$; $\underline{df} = 66$; $\underline{p} = < .01$). This difference disappears on the post-test with teachers and aides rating the children approximately the same. Both teachers and aides tend to rank the children with considerable consistency from the pre- to the post- measure; $r = .68$ for teachers and .74 for aides. Since the predominant change score was zero, it was not feasible to run a correlation between pre-test score and change score. In order to determine if any relationship exists between pre-test score and change score, a 3 x 3 contingency table was established consisting of a "gain," "no change," "loss" categories for the change score, and "above 3", "below 3" and "3" as pre-test ratings. The results of this analysis indicates no significant change on the part of the teacher ($X^2 = 3.14$; $\underline{df} = 4$; $\underline{p} = > .05$) but a significant change for the aides ($X^2 = 10.6$; $\underline{df} = 4$; $\underline{p} = < .05$). These data suggest then that the aides were more likely to shift their perceptions of the children and that their initial impressions of the children were less related to the change score.

Table 3 summarizes the results of the work vs. non-work responses to the items on the Stanford-Binet. These data are reported in terms of percentages of work responses to total responses and are summarized in terms of responses made to all of the items administered to a particular child and, in the bottom half of the table, the work responses made during the administration of the ceiling year tests for each individual child. Consistent with the findings of Hertzig et al, there is a greater tendency to emit work responses with respect to the performance items on the Stanford-Binet as opposed to the verbal items. It should also be noted that, as anticipated, there are somewhat fewer work responses during the ceiling year than on overall performance. Comparisons of pre- and post- performance indicate that there was no significant increase in the percentage of work responses for the verbal items of the Stanford-Binet, either in overall performance or with respect to ceiling year items. There was, however, a significant increase in the percentage of work responses to the performance items on both total performance and ceiling year items. Since the general trend for both verbal and performance items is in the direction of a greater emission of work responses, it seems reasonable to conclude that post-test performance on the Stanford-Binet is characterized by a greater tendency to attempt items rather than to avoid the demands of the situation. It is quite conceivable that the average gain in Binet I.Q. (6.1) is attributable to the greater number of task relevant responses made by the children on the post-test. In an effort to test this assumption, a correlation was run between Binet change score and the change score for work responses. This analysis resulted in a product moment correlation of zero indicating that a change in per cent work responses is not related to Binet change score. Our original formulation appears to have been oversimplified. Whereas we had anticipated that an increase in work responses would, even on a chance basis, lead to a gain in I.Q. score, a post hoc examination of the data indicates that this was not the case among the low pre-test Binet children; that is, their per cent work responses increased but they did not get more Binet items correct. Among the high pre-test Binet children, there was little gain in work responses, relative to the total group, but they made better use of their additional work responses than the low pre-Binet children. A similar case can be made for those children with average pre-test Binet's. Thus, it appears that an increase in work responses is more beneficial to children of average or above average ability than to children of low ability.

II. Social Behaviors

It will be recalled that four measures of social behavior were used: the Adaptive Behavior Rating Scale (ABRS), the Zigler Behavior Inventory, an adaptation of the Syracuse Scales of Social Relations with respect to personal attributes, and direct observations of the children's verbal behavior, conforming behavior, and attentional behavior.

With respect to the observations, we were not able, unfortunately, to secure enough data to warrant analyses in terms of pre-post measures. This occurred because we were only allotting ten minutes of observation time to a subject and, frequently, children were absent on the days that observations were being run.

Table 4 summarizes the three measures for which we have analyzed the data. All _t_ analyses are based on correlated means following standard procedures. It will be observed that the only significant changes with respect to our measures of social behavior occurred for the ABRS. These changes are in the direction of higher ratings on the post-tests and suggest that the children developed more adaptive behaviors, as defined by the teachers. Since this is the first time the ABRS has been administered anywhere, it is not possible to make comparative statements about this sample. However, to put the mean score in some perspective it might help to know that the maximum score possible is 310 (5 points is allotted to the "always" category and 1 point to the "never" category) and the minimun score possible is 62. If all of the children had received a scale score of 3 ("sometimes") on all of the items their score would have been 186. Thus, it appears that on the pre-test the teachers viewed the children as performing the behaviors on the average, somewhere between the categories of "sometimes" and "most of the time". The aides, it will be noted, rated the children lower; that is, somewhere between "once in a while" and"sometimes". On the post-test, the teachers ratings shifted in the direction of the category "most of the time" and this change is statistically significant. The aides, who show a greater degree of change, place the children, on the average, in the same relative position as the teachers. In terms of consistency of ratings over the intervention period, the product moment _r_ for teachers was .55 (_df_ = 31; $p < .01$) and for aides .31 which with 31 _df_ is not statistically significant. The correlation between the teachers' pre-test ratings and change score was .57 (_df_ = 31; $p < .01$) the similar correlation for aides was .31 degrees of freedom is not statistically significant. It is clear from these correlations that the teachers'

Table 4
ADAPTIVE BEHAVIOR

Inventory

	Teachers			Aides		
	M	SD	N	M	SD	N
Pre	196.7	33.6	33	183.4	31.2	33
Post	213.8	50.8	33	213.1	51.4	33
		t = 5.4**			t = 7.6**	

Zigler Behavior Inventory

	Teachers			Aides		
	M	SD	N	M	SD	N
Pre	138.5	20.6	34	139.7	21.9	32
Post	134.5	20.8	31	137.9	23.2	32
		t = 1.7			t = 1.2	

Personal Attributes

	Teachers		Aides	
	M	SD	M	SD
Pre	3.09	.91	3.00	1.29
Post	3.10	1.06	3.03	1.47
		t = < 1		t = < 1

** p = < .01

post-test score and consequent change score is more closely related to their initial rating of the children than is the case for the aides.

As will be noted in the next section of this report, the ABRS is significantly related to many of our cognitive measures suggesting that the instrument is measuring behavior of significance to the child's experience in the intervention program. As a measure of validity, we correlated the ABRS post-test scores with the Zigler Inventory post-test scores, resulting in a product moment r of .59 (df = 27; p = $<$ 01). (We did not run a similar correlation with the pre-test scores because of some pecularities observed on the pre-test scores for the Zigler.) It is clear that the Zigler and the ABRS share a significant portion of variance, but also, obviously, there is a good deal of unique variance between the two instruments. Unfortunately for the purposes of this report, we were unable to complete an item factor analysis of the ABRS, and are thus unable to report the behavioral categories reflected in our items. We anticipate completing this analysis soon making it possible to compare ABRS behavioral categories with those being identified by other investigators using the Zigler items. There are plans to readminister the ABRS by having the children's current teachers rate them as an estimate of between-teacher consistency. There are also plans to cross-validate the instrument using another nursery school sample.

The data summarized in Table 4 for the Zigler Behavior Inventory indicates a decline in ratings by both teachers and aides but the difference is not statistically significant. One possible reason for the lower post-test scores may be attributable to the difficulty reported by the teachers and aides in making the judgments required on the Zigler so early (first week) in the intervention program. They also reported difficulty in understanding the more "global" items. Despite these comments it should be noted that the correlation between the pre-test ratings and post-test ratings for the teachers was .80 ($df = 31$; $p = < .01$) and a similar correlation for the aides was .72 ($df = 31$; $p < .01$). Among our teachers and aides, therefore, there was a tendency to rate the children slightly lower on the post-test with relatively little shift in the rank ordering of the children.

The third set of data in Table 4 relates to the 'personal' attributes ratings using the Syracuse Scales of Social Relations. 'Personal' attributes, as defined here, refers to the general social attractiveness of the children and not to any particular attribute of social behavior. Recall that these ratings are made on a 5-point scale where the raters defined the two end points and the mid point in terms of all the kindergarten children they had ever known. The resulting data reflect the relative position of our sample of children with respect to the reference populations defined by the teachers and aides. Examination of the means for both teachers and aides indicates that the average child in the sample was placed in the almost exact middle of the scale. This is in contrast to the very much lower ratings given to the children on the 'academic' attributes scaling. This discrepancy suggests that despite the low academic attributes of the children, their personal attributes are seen as average. As shown in Table 4, there were no statistically significant changes between pre- and post- ratings for either teachers or aides on the 'personal' attributes scale. The correlation between the initial rating and the post-rating was .47 ($df = 31$; $p < .05$) for teachers and .47 ($df = 31$; $p < .05$) for aides on the personal attributes scale. In contrast, the identical correlations for the academic scale ratings were substantially higher (.68 and .74 for teachers and aides respectively). Thus, for both groups of raters there is a greater relationship between pre- and post-test scores for the academic attributes as opposed to the personal attributes.

III. Input Measures

There were two measures of stimulus input used in this study: 1. the inventory of home stimulation, and 2. observations of the teacher and aide behavior in the classroom.

Table 5 summarizes the means and SD's for both measures. Because of the relatively little interaction occurring between the children and the teachers and aides it was decided to pool their praise and blame statements to each of the children. Despite the few observations that were recorded, it can be seen that the teachers and aides were tending to use more blame in their approach to the children than praise. The difference between the praise and blame received by the children is statistically significant (t = 2.2; df = 31; p = $<$.05). It should also be noted that the variation in blame received is greater than for praise received. This occurred because three or four of the children received a comparatively large number of the blame contacts from the teachers. These results are not unlike those reported by others (Meyer and Thompson, 1956) and suggest that some children learn very early to expect disapproval from school personnel.

The mean and standard deviation for the STIM is also reported in Table 5. Since there are no norms available for this version of the STIM, it is not possible to comment in any intelligent way about the meaning of the mean and the standard deviation. Instead, we will use this measure in the next section of this report that discusses the inter-relationships among the measures.

IV Inter-relationships of Measures

This section of the report is concerned with the inter-correlations among the variables used in evaluating this intervention program. In view of the fact that the data were not available for analysis until mid-September, it will only be

Table 5
Means and SDs for
Stimulus Input Measures

	Praise (Teacher and Aide)	Blame (Teacher and Aide)
M	2.03	3.44
SD	1.68	3.70

STIM
M = 46.4
SD = 10.4

possible to report a few of our findings. As further analyses become available, they will be included in subsequent quarterly reports and other written materials. One other comment is necessary before examining the data. We are only too well aware that correlational data are subject to a large number of variables that may have occurred fortuitously as a result of sampling peculiarities among the teachers and/or among the children. In reading this material, much of which strikes us as exciting, it should always be remembered that the data are derived from one sample of children whose totality of characteristics are largely unknown to us.

Recall that the Draw-A-Line Slowly and Walk-A-Line Slowly Tests were designed to measure "impulsivity". It was assumed that impulsive children are less likely to attend to the stimulus characteristics in their environments and are more likely to respond to specific characteristics or components of the stimulus situations without adequate analyses of the total stimulus situation. If the DAL and WAL reflect the described cognitive style there should be a significant relationship between these measures and the Binet. Impulsivity should also be related to the adequacy of social behavior in the classroom. Thus, the impulsive child may talk too much, may become involved in disputes with other children more frequently, or he may run down hallways instead of walking.

The first set of correlations focus on the relationship between the impulsivity measures and performance on the Stanford-Binet. The correlations involving the DAL are based on the "rate" measure; that is, length of line/time. (Correlations were run between absolute time elapsed on the DAL and Stanford-Binet and they are the same magnitude or greater, than those to be reported now.) With this measure, a high score signifies a greater degree of impulsivity than a low score. Thus, if impulsivity is related to Binet performance, we would expect negative correlations.

Scores on the WAL are direct time measures, and they would be expected to correlate positively with Binet performance. This discussion also includes an analysis of both the error and latency scores on the perceptual speed test. A negative correlation between errors on the PST and Binet would be anticipated, whereas a positive correlation between latency on the PST and Binet is anticipated.

Table 6 provides the relevant correlations showing the pre-test correlations in the upper part and the post-test correlations in the lower part of the Table. Examination of Table 6 indicates that both the WAL and the DAL are significantly related to performance on the Stanford-Binet, for both pre- and post-test. The PST measures, for the pre-test, are not significantly related to any of the other measures. For the post-test assessment, however, the PST measures are significantly related. Indeed, it should be noted that the correlations for the post-test measures are generally higher than those for the pre-test measures.

One obvious question that can be raised concerning the correlations between the DAL and Binet and the WAL and Binet is the degree to which performance on the impulsivity measures reflects the ability to understand the instructions. (This is an interesting question in that it implies, from our data at least, that understanding instructions accounts for something like 36% of the variance on Stanford-Binet). The question can be answered, at least partially, from procedures used in this study with the WAL. Prior to giving the children instructions about going "slow" and going "fast", we asked the children to walk the line without any instructions. The correlation between performance with no instructions and pre-test Binet was .43 and when the same procedure was repeated on the post-test the correlation with post-test Binet was .51 (both correlations are statistically significant at the .05 level). Of greater interest is the correlation of .81 between the "no instruction"

Table 6
Inter-Correlations of the DAL, WAL, PST and Stanford-Binet for Pre and Post-Test (Decimal points have been deleted).

	1	2	3	4	5
			PST		
	WAL	DAL	errors	latency	S-B
1		-40*	-10	-06	44*
2	-72**		17	-23	-45*
3	-39*	44*		-29	-26
4	51**	-57**	-10		-11
5	60**	-56**	-43*	43*	

* p => .05
** p => .01

condition and the instruction to walk "slowly", but only a correlation of .38, which is not statistically significant, between the "no instruction" condition and the instruction to walk "fast". These data suggest that children who are impulsive under the "no instruction" condition are unable to conform to the "slow" condition but that the more controlled children under the "no instruction" condition can conform to the "fast" instructions. We have already noted that there is a statistically significant difference in performance under the "slow" instructions as opposed to the "fast" instruction condition. Finally, it should be noted that all the correlations involving the WAL and the DAL in the "fast" condition and the Stanford-Binet measure are essentially zero. Thus, it would appear that the correlations for the DAL and WAL with the Stanford-Binet are not a function of the childrens' ability to understand the instructions; namely, the distinction between "slow" and "fast". Rather, it would appear that those children who are able to conform to the "slow" instruction are able to utilize environmental information than more impulsive children. This latter assertion receives some support from the fact that the correlation between the WAL pre-test and change score on the Binet is .56 which indicates a significant relationship between I.Q. point gain and motor control.

One of the interesting, and unexpected, findings in this study were the generally higher correlations among the post-test measures; particularly where latency measures were involved. It was found, for example, that correlations between pre- and post-test latency measures were very low, and in some instances were even negative. The pre-test latency measures, however, generally do not correlate with any of our other measures but the post-test latency measures correlated significantly. Our interpretation of these results is that, perhaps as a result of the intervention experience, or as a result of their pre-test experiences, the within subject variability over tests was sharply reduced; that is, the post-test measures reflect the child's general test taking style. The alternative explanation that between subject variability increased from the pre- to the post-tests is not supported by the data. It would appear, therefore, that our post-test measures provide a more accurate picture of the child's behavioral capability, especially with respect to latency measures. These data also suggest that too early administration of a pre-test battery may not provide the best estimate of the child's range of capabilities.

One final comment on the DAL and WAL tests. Our results are in general agreement with those reported by Maccoby *et al*. (1965) whose subjects were quite different from this sample, at least in terms of average intellectual ability. The results are not in agreement, however, with Hess *et al*. (1966) and Banta (1967) who failed to find significant relationships between the DAL and Binet performance. Obviously, more research is needed to define the correlates of impulsivity and, more importantly, to determine under highly controlled conditions the behavioral consequences of impulsivity.

Table 7 summarizes the correlations between the DAL, WAL, Binet, and the ABRS. Our concern here was with the degree to which the teacher's perception of the adaptive behaviors of the children were related to our impulsivity measures and Stanford-Binet performance. These correlations indicate that both the teacher's and aide's perceptions of the adaptive behaviors of the children are significantly related to the children's performance on the Stanford-Binet. With one exception, the impulsivity measures are also significantly related to the ABRS, but only for the pre-test scores. It is quite possible that during the six-week intervention program the more maladaptive social behaviors of the more impulsive children were

Table 7
Intercorrelations of ABRS with DAL, WAL, and
Stanford-Binet for Pre- and Post-Tests
(Decimals have been deleted)

	Teacher		Aides	
	Pre ABRS	Post ABRS	Pre ABRS	Post ABRS
Binet	43*	43*	46*	52*
WAL	22	-05	44*	14
DAL	-58**	-30	-55**	-29

* p = >.05
** p = >.01

extinguished, and that they learned to make the more elementary responses considered important by the teachers and the aides. It is also clear, however, that as perceived by the teachers and aides the brightest children tend to make the most socially adaptive responses. Incidentally, it should be noted that none of the cognitive measures were significantly correlated with change score on the ABRS.

V. The Arithmetic Project

In addition to the regular classroom porgram, the research staff conducted two special programs; 1. a Montessori Program, and 2. a Bereiter type program focusing on arithmetic only. The Montessori class was conducted by Mrs. Jacquline Morrow Massari who had previous experience with a Montessori class but with no formal training in the techniques. The Bereiter group was conducted by Miss Gwen Simpson, under the direction of Dr. Vernon Hall.

Two Montessori groups of five each were established by randomly selecting the 10 children from among the 33 enrolled in the program. The Bereiter group consisted of five randomly selected children from the sample. A total of eight randomly selected children served as control subjects.

Since the Bereiter group only worked on arithmetic, comparisons among the groups were based only on arithmetic achievement (undoubtedly other perhaps more important variables should have been assessed but time did not allow for anything more. Besides, there is considerable confounding because of the design used. Indeed, the entire project must be considered as exploratory.) All groups were administered an arithmetic achievement test developed for this specific purpose. A pre-test post-test design was used.

The children in the three experimental groups, two Montessori and one Bereiter, went to their particular rooms for one-half hour each day. The control group remained in their own classrooms. Montessori materials were available to the children and no special emphasis was arithmetical operations or concepts. The Bereiter group followed the program used by Bereiter.

Although statistical analysis are not completed, there may be some interest in how the children performed on the test (see Appendix A for a copy of the test). Table 8 summarizes the means and SDs for the pre- and post-tests for each group. These data indicate that the largest gain occurred in the Control Group and is statistically significant ($\underline{t} = 4.6$; $\underline{df} = 7$; $p < .01$). The average gain for the pooled Montessori groups is not statistically significant ($\underline{t} = 1.8$; $\underline{df} = 8$; $p > .05$) nor is the gain for the Bereiter group statistically significant ($\underline{t} = 1.6$; $\underline{df} = 4$; $p > .05$).

Perhaps the most significant aspect of the data in Table 8 is the substantial variation in performance on the test. Examination of pre-test scores for the entire group shows a range of 0-23 and a post-test range of 4-29. These data suggest that some preschool children are capable of handling some fairly sophisticated arithmetical concepts and operations.

Discussion

Although all of the comparative and inter-correlational analyses are not completed, several implications would seem to emerge from the available data.

Considering first, the various measures of cognitive ability, it is clear that in our sample a meaningful proportion of the variance in Stanford-Binet performance is related to performance on the "impulsivity measures". In particular, we are impressed with the relationships between the DAL, the WAL, and Binet performance, because neither the DAL or the WAL would appear, on the surface, to be a measure of what is commonly called "g". This is particularly clear with respect to the WAL test where no relationship was found between walking a line fast and Binet performance but where correlations of .50 were found between walking a line "slow" and Binet and walking a line without any particular instructions and Binet. The "no instruction" condition would seem to indicate how children respond to a variety of demands in a more naturalistic setting. Thus a hypothesis that impulsive children respond to a broad variety of stimuli, both relevant and irrelevant, in their environment without integrating them or establishing appropriate response patterns to these stimuli seems to be indicated. Our data, along with those reported by Maccoby _et_ _al_. (1965), suggest that the deleterious effects of impulsivity are as apparent among children of superior intellectual ability as among those with low normal ability.

Perhaps the most significant aspect of our findings with the DAL and WAL is the fact that no relationship was found between "fast" condition and the Stanford-Binet. This finding coupled with the fact that the difference in time under the "slow" and "fast" condition was statistically significant suggests that the correlation between the "slow" conditions and the Binet is not an artifact of understanding the instructions.

Our findings with respect to the children's responses to the cognitive demands of the Stanford-Binet were somewhat disappointing in that increases in per cent work responses were not especially related to increases in Stanford-Binet I.Q. Apparently, our hypothesis was only accurate for those children in the middle I.Q. range of our distribution, but not for those children on either extreme. This raises certain interesting questions. One possibility is that improved motivation leads to a greater tendency to give relevant responses to the Stanford-Binet. This assumption was supported by our data but, as already noted, increased work responses were not demonstrably related to I.Q. gain. In this context it should be kept in mind that the designation of a "work" response is not determined by whether or not the answer given to the question is correct. Thus a child who increases in "work" responses is now putting out a necessary but not sufficient response in order to get a Binet item correct. It may be hypothesized then that in order to get an item correct a child must both have a set to work at answering and also have acquired the necessary content. It would be expected that a general "non-work" mode of response would mitigate against learning throughout the child's development. A change in tendency to work, therefore, would not result in an increased Binet score unless or until the work tendencies had operated in the acquisition of sufficient content to enable the child to have the "right answer" response in his repetoire. A change in "work" tendency may then be seen as a very real improvement with which other cognitive inputs will interact.

Another possibility, more difficult to examine empirically, is that initially poorly motivated children in fact emit work responses but make little or no effort with respect to the accuracy of their response. Indeed, examiners have been known to report that children sometimes seem to make errors on purpose. If this were the case, our work-non-work categories would be unable to detect such subtle changes in behavior.

Several aspects of the results of our examination of the teachers' perceptions of the children merit comment here. As just about everyone would expect, on the basis of studies such as Rosenthal's (1966), the teachers perceptions of children's social adaptiveness is significantly and positively correlated with intelligence. This finding occurs despite the substantial precautions taken to avoid communicating to the teachers about the intellectual abilities of the children. The issue may not be as simple as the notion that teachers like bright children only because of their intellectual competence. Our data suggest, for example, a tendency for the less able children to display a greater degree of impulsivity. Many of the adaptive behaviors specified by the teachers require, for example, substantial motoric control. In other words, it is important to teachers that children stand quietly in line or sit quietly in their seats while they are giving instructions. The high impulsive child is less likely to perform this task, thus becoming a source of irritation to the teacher. Given a broad range of intellectual abilities, it would be surprising indeed to find no relationship between intellectual and social abilities. These data suggest that it may be well to formulate our theoretical models of child socialization in terms of cognitive structures. It could be argued that children who can grasp and act upon conceptualizations will be in a better position to understand and accept the demands made by their teachers. It also follows that the more impulsive child, who may find his relationships with teachers more difficult even with mature cognitive structure, is more likely not to have such structures available to him. Thus, this child is placed in a much more difficult situation as suggested by the low ratings they tend to receive from their teacher. We do not mean to imply that all of the variance in teacher attitudes in children can be accounted for in terms of variation in cognitive structure. Certainly there are variations in the social adequacy of children's behavior even where there is homogeneity with respect to intellectual ability and cognitive style of functioning. Such variations undoubtedly result from important influences shaping the child's social behaviors, such as parent-child and teacher-child relationships. Our data provide only the most minimal cues as to what happens to the less positively perceived children, namely, that they tend to receive a substantial proportion of the teacher's blame statements without necessarily a concomitant increase in praise for more controlled behavior. Manipulation of unapproved behaviors is apparently attempted almost entirely by punishment albiet verbal and not particularly severe.

A final observation. Throughout the arduous business of data analysis, one could not help but feel overwhelmed by the variation among the children on each of the measures and the concomitant gain scores. This feeling of concern about the variation is certainly not unique to our research group, but one can't help wondering why classroom programs continue to have a uniformity of procedures over all children. Perhaps future research efforts should concentrate on identifying the salient dimensions of variability (multi-variable statistical techniques are now available) and, then concentrate on variations in programming for individual children.

Table 8
Means and SDs for Arithmetic Test

	Pre-Test		Post-Test		Gain
	M	SD	M	SD	
Montessori I	10.4	7.6	13.2	7.3	2.8
Montessori 2	7.7	2.5	9.7	2.8	2.0
Bereiter	9.2	8.2	12.4	6.8	3.2
Control	9.9	7.2	15.2	9.9	5.3
Total	9.4	6.5	13.1	7.5	

References

Banta, T. Tests for the evaluation of early childhood education. The Cincinnati Autonomy Test Battery. Mimeo, 1967.

Caldwell, Bettye, Heider, Jaye, Kaplan, Barton. The Inventory of Home Stimulation Delivered at APA, New York, September 1966. Mimeo, 1966.

Flanagan, J. A proposed procedure for increasing the efficiency of objective tests. Journal of Educational Psychology, 1937, 28, 17-21.

Gardner, E. and Thompson, G.G. Social relations and morale in small groups. N.Y.: Appleton-Century-Crofts, 1956.

Hertzig, Margaret E., Birch, Herbert G., Thomas Alexander, and Mendez, Olga Aran. Class and Ethnic Differences in the responsiveness of pre-school children to cognitive demands. Mimeo, 1967.

Hess, R., Kramer, R. Slaughter, Diane, Torney, Judith,Bury, Carla, and Hall, Ethel. Techniques for assessing cognitive and social abilities of children and parents in Project Head Start. Mimeo, University of Chicago, 1966.

Kagan, J., Moss, H. and Siegel, I. Psychological significance of styles of conceptualization. Monographs Society Research Child Development, 1963, 28 (#86).

Maccoby, Eleanor, Dawley, Edith M., Hagen, J.W., and Degerman, R. Activity level and intellectual functioning in normal pre-school children. Child Development 1965, 36, 761-770.

Meyer, W.J. and Thompson, G.G. Sex differences in the distribution of teacher approval and disapproval among sixth-grade children. Journal Educational Psychology, 1956, 47, 385-396.

Zigler, E. Personal Communication

VARIABLES AFFECTING THE PERFORMANCE OF YOUNG CHILDREN ON A LETTER DISCRIMINATION TASK

By Vernon C. Hall, Edward Caldwell and Gwen Simpson

The purpose of the present research was to determine the influence of several variables on the performance of a letter discrimination task by young children. It was felt that this would help ascertain the proper conditions under which readiness tests should be administered in order to more accurately decide who needs training. In addition, it should provide new information as to what age children typically exhibit the ability to correctly identify d, b, g, and n.

Depending on the specific goal desired, there are several strategies which can be used in setting up the proper conditions under which a standardized test is given to young children. In the area of aptitude testing the major requirement is usually that the administration be as uniform as possible across subjects, since we are interested in individual differences. This is typically called a norm referenced test. If we assume that all children are trying equally hard the fact that different conditions (i.e., different instructions) might improve the performance of all subjects is not a problem since any individual child will maintain his same relative position.

When we give tests designed to discover what the child actually knows so we can determine what to teach next, we are interested in the ability of the child to perform a specific task. In this situation it is extremely important that we structure the situation in such a way that the performance of the child is as close as possible to his true capacity. This is typically called a criterion referenced test. It is the contention of the present writers that in typical studies designed to gather this kind of information not enough attention has been paid to this problem. The present study was designed to investigate the influence of five variables (initial instructions, motivation, warmup condition, letter size and feedback) on the performance of a letter discrimination task by kindergarteners.

The strategy was to replicate a widely quoted study of letter discrimination by Davidson (1935) which reported age norms on the acquisition of the ability to discriminate the letters d, b, q, and n (along with other letters). In addition each of the five conditions mentioned above would be varied to determine what effect they had on performance. Therefore, in order to explain the rationale behind the actual procedures used in defining the above conditions it is necessary to briefly discuss the procedures and findings of the original Davidson study. This might best be done by quoting Davidson. "The Letter Perception Test devised for this experiment was in the nature of a cancellation test. A certain letter was printed in large black type at the left side of a sheet of paper. After it a 'box' was drawn containing four rows of ten letters each. Among these letters were five identical to the given letter, five of the letter most likely to be confused with it (the mirrored opposite in the case of d, q, and b), several each of other possibly confusing letters including several tall letters in case 'tallness' was an important factor in letter recognition, while the remainder were randomly selected letters. The children were instructed to look at the given letter, then find every one exactly like it in the box and mark it when they found it. There was one practice exercise (p. 459)."

When scoring the results the author explains that since some children did not mark all five of the letters "A child has been credited as having confused b and d whether he marked one b or five b's." The results indicated that a mental age of 6 was required for up down confusions and 7 and one half years was needed before left right discrimination were consistently correct. This is consistent with Wilson and Fleming (1938) who found confusion persisting to the age of 8 or 9 years.

One strategy in taking a test is to mark all answers which might possibly be correct thinking that some will be right. On the other hand when the S is aware that some kind of correction procedure is being used (i.e., if one is wrong all are wrong) then the correct strategy would be to be more careful. Thus, in the present study one of the deviations from the original Davidson procedure was to instruct the children that one incorrect letter on a single row of letters made the entire row incorrect.

With regard to motivation it was speculated by the experimenters that the group standardized testing situation in which there are no real incentives to perform well (i.e., the fact that it is an outside administrator telling them they won't be graded for their performance) may lead some children to not try very hard or even attend. In a discrimination task such as the one used, close attention is particularly important. Thus it was hypothesized that a tangible reward (candy) contingent on correct responses would be influencial in improving performance.

The brief mention of a warmup condition in the original study left us with somewhat of a problem. The experimenters felt this was an extremely crucial variable because if the child's definition of same is different than that of the experimenter's a wrong answer will occur. In the present experiment we designed the warmup to either measure a very global meaning of the word same and different (i.e., triangles and circles) and one in which orientation (right and left, up and down) of letter like forms was the distinguishing feature.

It has often been mentioned by reading experts that the size of the letter is important for young children. For instance, Bloomfield and Barnhart (1961) tell us "Under no circumstances should you start to teach your child to read until he has the necessary muscular skills to see the small distinctions between the letters of the alphabet. You can tell he has the necessary muscular coordination when you notice that he can button his clothes or see and pick up a pin on the floor (p.3)." Although the actual size of letters used by Davidson were not mentioned we had planned to use pica size type. When the principal of an elementary school listened to our research plan, he mentioned that this was much too small for young children and that all elementary schools had typewriters with primary size type. Therefore, we decided to use letter size as a variable. Half of the children received the task using pica type and half using primary type.

Finally since the group receiving candy was to also obtain information as to how they did on the task (in order to determine how many candies they were to be rewarded) it was decided to see how important this feedback was for future performance. It was also frowned on by some that not all children receiving candy. Therefore on the day following the original presentation all children were re-administered the same task and everyone received candy for doing the task regardless of how well he did.

It should be clear that the experimenters did not view any of these procedures as training in the traditional sense because at no time was any overt instructional information given. It has already been shown that kindergarten children can be trained to improve on similar types of discrimination tasks (i.e., Jeffery, 1958, Hendrickson and Muehl, 1962). The experimenters are trying instead to show that for many children under typical standardized conditions the difference between letters and non letters is already perceived but they don't know how the adult defines same and different. If this is true, then many hours of educational time may be wasted in trying to teach children something they already know.

Subjects: Subjects were 80 kindergarten students attending the Jamesville-Dewitt elementary school, Jamesville, New York. Their chronological ages ranged from 5 years 5 months to 6 years 4 months with 60% being below 6 years of age. They were randomly assigned to one of five treatment groups as they entered the experimental room.

Methods and Procedure: Each child was brought into the experimental room and the appropriate instructions were read to him. After the instructions were read the experimenter left the presence of the S and returned only if the S had any questions or when the S had completed the task. The experimenters at no time deviated from the written script except to reiterate instructions or encourage the subject to keep going. The warm up (Figure 1) required the subject to complete two lines and then he was shown how well he had done.

The actual design consisted of two letter sizes and the five following treatments: 1) Replication of the Davidson study but administered individually using circles and triangles for demonstration; 2) Same as condition 1 except the S is instructed that if he circles one letter incorrectly the entire line is incorrect; 3) Same as conditions 1 and 2 except that the S was presented with one M&M candy for each correct answer but if one is incorrect then the S loses all candy won for that row; 4) Same as condition 3 except that symbols with relevant cues for the final task were used in the warmup; 5) Same as condition 4 except that M&M candy was used as in condition 3.

Upon the completion of the task Ss in conditions 1, 2 and 4 were sent back to their classroom. Ss in conditions 3 and 5 were shown which items were right and which were incorrect and rewarded with their candy. No mention was made of why individual items were correct or incorrect. The next day all Ss completed the same task and were given four M&M candies.

Results: Figure 2 shows the percentage of Ss for each condition which got 100 per cent right for each letter along with the appropriate percentages for the original Davidson sample. A 2 (letter size) by 5 (treatment) analysis of variance on the first test was computed using percentage correct. The performance on letter size was almost identical (F = .34, 1/70 df). Although the treatment effect was only significant at the .08 level (F = 2.27, 4/70 df) examination of Figure 2 shows that Ss in condition 5 outperformed all of the other conditions and at a level equal to that of children with a C. A. of 7-6 to 7-11 in Davidson's sample on letters d and n, better than 7 year olds on the b and out performed the 6-6 year olds on the q. There were no significant interactions.

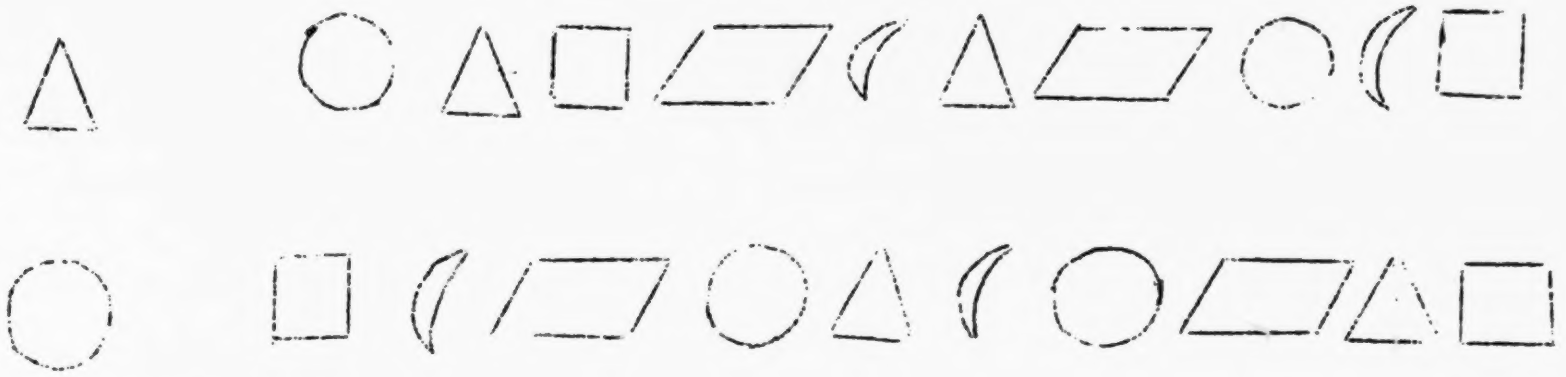
The second test clearly exhibits different results. Conditions 3, 4 and 5 all improved on every letter. A 2 (letter size) by 4 (individual letter) repeated measures by 5 (treatment) analysis of covariance removing performance on the first test was computed. Again the letter size performance was almost identical but the letter (F = 40.96, 3/240 df) and treatment effects (F = 2.98, 4/70 df) were significant. In addition, a Scheffe analysis indicated that treatments 4 and 5 were significantly better than 2 and 3, treatment 5 was significantly better than 3, and 3 and 5 were significantly better than 1, 2, and 3. Finally the children in treatment 5 outperformed all ages used by Davidson (he has children up to 8 years of age).

Discussion: There is little doubt that the results of this experiment indicate that letter size is not an important factor with kindergarteners. In all cases performance on both letter sizes was almost identical. In addition the instruction given to the child with regard to scoring procedures had little effect. Although it may be that these differences in instruction were not powerful enough the researchers tend to think that kindergarteners are not yet sophisticated enough test takers to apply different strategies when a rights minus wrongs procedure is used.

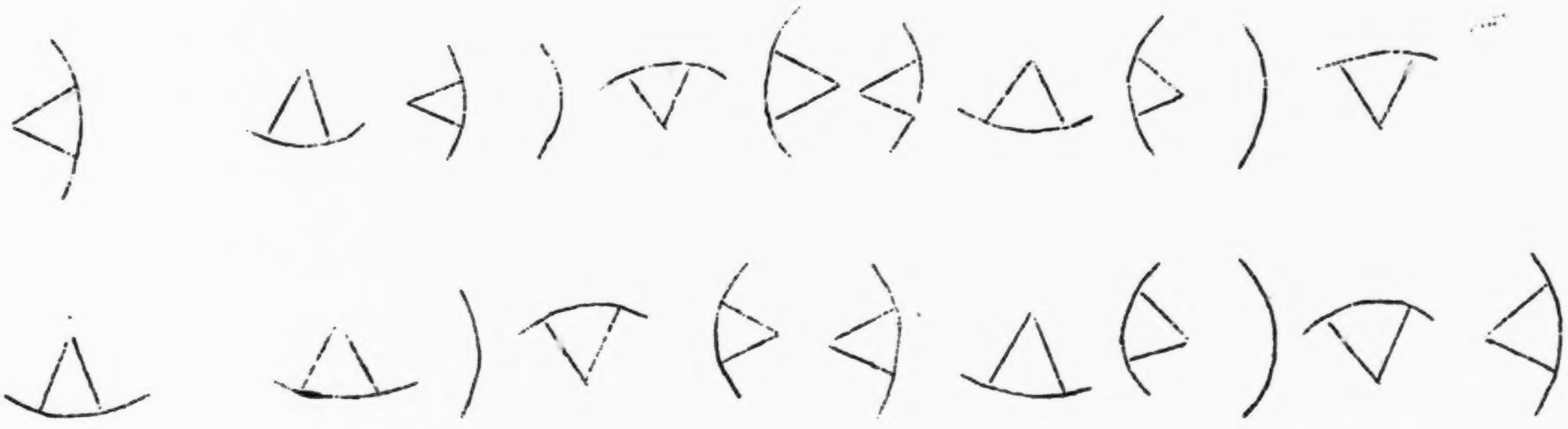
The picture with motivation, warmup and feedback isn't that clear cut. It is of course, obvious that warmup condition is more important than the motivation factor in the sense that when children are unaware of what the correct answer is, no amount of trying will help. The authors' interpretation is that the most important variable was knowing what the experimenter meant by "same" and "different." The fact that the feedback was so effective indicates that a longer warmup or one in which the symbols had a closer resemblance to the letters would have been even more effective. The authors would see no real disadvantage in having the warmup and the task identical.

The real importance of the study is that it indicates the performance of a young child in this kind of task is highly dependent on the testing conditions. Regardless of statistical significance it is evident that there were probably a number of individual children who were misclassified as being incapable of discriminating letters by Davidson,

The authors believe they have demonstrated that it would be well worth the time of any test constructor interested in validating criterion referenced tests to do some preliminary investigations into the proper conditions under which the test should be administered.



Warmup stimuli for groups 1, 2, and 3.



Warmup stimuli for groups 4 and 5.

Figure 1

Group		N	d	n	q	b
Davidson Study						
	Ages 5-6 through 5-11	34	10.0	52.5	10.0	12.5
	Ages 6-0 through 6-6	40	26.5	76.5	32.3	41.2
Present study first test						
Condition Number	1	16	12.5	50.	12.5	6.25
	2	16	25.	50.	18.75	25.0
	3	16	12.5	75.	12.5	12.5
	4	16	18.75	68.75	12.5	37.5
	5	16	50.	87.5	37.5	43.75
Present study second test						
Condition Number	1	16	12.5	50.	18.75	18.75
	2	16	25.	62.5	18.75	31.25
	3	16	50.	87.5	25.	37.5
	4	16	43.75	75.	43.75	43.75
	5	16	75.	93.75	68.75	68.75

Figure 2. Percentage of Subjects Selecting d, n, q, and b without Error.

References

Davidson, H.P. A study of confusing letters B, D, P, and Q. *Journal of Genetic Psychology*, 1935, 47, 458-468.

Hendrickson, L. and Muehl, S. The effect of attention and motor response pre-training on learning to discriminate B and D in kindergarten children. *Journal of Educational Psychology*, 1962, 53, No. 5, 236-241.

Jeffrey, W.E. Variables in early discrimination learning: 1. Motor responses in the training of a left-right discrimination. *Child Development*, 29, No. 2, 269-275.

Wilson, F.T. and Fleming, C.W. Reversal in reading and writing made by pupils in the kindergarten and primary grades. *Journal of Genetic Psychology*, 1938, 53, 3-31.

APPENDIX A

Adaptive Behaviors

Child's Name ____________________

Pre. ___________

Post. ___________

Date ___________

Teacher ____________________

Please indicate as accurately as possible how this child behaves by marking one of the five responses for each item. Base your response to every item on your personal observation and experience with the child.

	Never	Once in a while	Some times	Most of the time	Always
1. Is able to go to the toilet by himself..					
2. Can get from one room to another in the building himself......................					
3. Can use eating utensils to feed himself					
4. If sick or hurt can report to parent or teacher and tell what hurts............					
5. Will share play equipment when requested by teacher.........................					
6. Will share appropriate equipment when requested by another child............					
7. Reports infringments on own equipment, food etc. by another child to teacher or gets it back peacefully............					
8. Uses more than one color or material when making, painting or decorating pot or bowl................................					
9. Knows his bus when it is time to go home					
10. Knows where he lives.................					
11. Is orderly in line....................					
12. Knows name of teacher.................					
13. Knows class mates' names...............					
14. Accepts consequences of own behavior i.e. does not blame others for own accident................................					
15. Responds to questions about pictures, etc. with more than grunts or shrugs. i.e. can describe things...............					
16. Comments spontaneously about pictures, exhibits etc..........................					
17. Describes feelings of like or dislike about things..........................					
18. Attends to a sit-still activity such as story for at least 10 minutes..........					
19. Attends to a work activity such as paint or clay for ½ hour....................					
20. Cleans up after himself and helps clean up area................................					
21. Recognizes a photograph of himself.....					

	Never	Once in a while	Some times	Most of the time	Always
22. Can name the primary colors					
23. Notices nature of and changes in properties of objects i.e. color mixing, form of clay, growing plants, etc.					
24. Knows his own class room					
25. Can hold and control a pencil or brush					
26. Can stay inside lines when coloring an outline form					
27. Follows verbal directions					
28. Waits for directions before rushing in					
29. Does not grab					
30. Answers when called on					
31. Waits in line and stays in his place					
32. Obeys safety rules					
33. Remembers safety rules					
34. Keeps himself relatively clean, and gets cleaned up if he gets dirty					
35. Knows names of kitchen utensils and how used					
36. Knows shop tools and what used for					
37. Knows TV commercials and characters					
38. Asks questions if doesn't understand words or directions					
39. Can complete an idea if teacher stops in middle of sentence					
40. Uses I, we, he when speaking					
41. Makes verbal relations between what is happening and other incidents in or out of school					
42. Can give verbal description and reasons for his behavior					
43. Pretends enthusiastically					
44. Changes verbal mistakes in grammar when instructed					
45. Can cut small shapes approximately 2 inches with scissors					
46. Can copy simple geometric figures, circle, triangle, square etc					
47. Can draw simple designs and some letters with crayon					
48. Handles his body in a coordinated way.					
49. Knows male from female					
50. Will not respond when he knows the answer					
51. Swears					
52. Parrots verbal behavior of other children inappropriately					
53. Clings to teacher					

	Never	Once in a while	Some times	Most of the time	Always
54. Interferes with other children's work or activity..........................					
55. Demands attention when teacher is busy					
56. Whines............................					
57. Physically weak and uncoordinated.....					
58. Crys without sufficient reason........					
59. Sucks thumb, picks at self, always fighting..............................					
60. Stays by himself and doesn't mix with other children.......................					
61. Shoves or hits other children.........					
62. Will not attempt new activity.........					

STIM Guide to Interviewers

The questions in this inventory consist of two general types. First those which you can check from your observation during the interview and those about which you will have to inquire. Generally inquiry items are those with an asterisk. Within the second category there are some items which may easily be asked directly as they stand in the questionnaire and those for which notations are made in this guide for eliciting the answer.

Each question is listed by number and is followed either by the notation Observe; Ask, which means the question may be asked as it stands; or by a notation to serve as a guide to your phrasing of the question. Please attempt to get an answer which may be checked yes or no for all items, and fill in the face sheet completely.

Items

Face Sheet: Ask

I. Adult Contact.

1. Do you work? (If mother is working more than half time she is not the principal caretaker).
2. When you are away from the house do you have a regular sitter or sitters?
3. Observe.
4. Where is (child's name) usually, when mother is doing housework? Does she talk to him while she is working?
5. Observe.
6. Does _____ know when to expect meals, his nap?
7. When does your husband get home? What does he usually do with the child? Games, dressing, sitting when mother is away etc.

II. Developmental and Vocal

8. Is more than one language spoken in the home?
9. Do you read to _____ and/or talk about picture books together? About how often?
10. Observe.
11. Observe.
12. Observe.
13. Observe.
14. Observe.
15. Observe.
16. Ask
17. When _____ was younger, you probably encouraged him to do things like walk and talk,what things like that do you do now? e.g., tying own shoes, table manners, concentrating on a play actively, generally any of the developmental advances appropriate to this age.
18. Ask
19. When _____ gets a new toy how do you like him to play with it? Do you show how it works?
20. When _____ is not using his toys where are they kept? Are some activities introduced on purpose by mother?
21. Ask
22. Observe.

23. Ask
24. Ask
25. If _____ wants you, do you go right away or say wait?
26. Same as 25 except child's call is more urgent.

III. Emotional Climate
27. Observe.
28. Observe.
29. Observe.
30. Observe.
31. Observe.
32. Ask
33. Observe.
34. Observe.
35. Observe.
36. Observe.

IV. Avoidance of Restriction
37. Observe.
38. If _____ spills his food or drink, what do you do?
39. Observe.
40. This item is Yes if 43 is Yes, otherwise ask what he plays at indoors.

V. Experience
41. Ask
42. Ask
43. Ask
44. Ask
45. Ask
46. Ask
47. Ask
48. Ask if he has regular checkup. Code as No if only goes when sick.
49. How much of the days is the television on. Code No if on most of the day.

VI. Physical Environment
50. Observe.
51. Observe.
52. Observe.
53. Observe.
54. Observe.
55. Observe.
56. Observe.
57. Observe.
58. Observe.
59. Observe.

VII. Play Materials

60. Observe or ask for each class of toy for items 60-68.

Write any additional comments or qualifications necessary on sheet at end of questionnaire.

Math, Pre and Post Test

E's name ____________________ Child's name ____________________

1. Count to 20 (Record errors as they occur)

2. (Place 20 checkers on the table, in 2 rows of 10 each.) Count the checkers like (demonstrate counting 3 checkers, moving each one slightly as you count aloud). Now, you count all of the checkers and tell me how many there are.

3. I will show you some numbers. Tell me the name of each number. (Show the number cards one by one.)
   (5)

   (13)

   (0)

4. Here are some checkers (place 10 checkers on the table, in 2 rows of 5 each) and here are some cards with numbers on them (spread number cards 1-10 on the table). Chose the card that tells how many checkers are on the table. (Repeat with groups of 2 and 7 checkers)

   (10)

   (2)

   (7)

5. I'm going to clap my hands. Tell me how many times I clap my hands. (Clap first 2 times, then 5 times.)
   (2)

   (5)

6. a. How many are 3 plus 2?

   b. How many are 2 plus 3?

   c. How many are 8 and 0 (zero)?

   d. How many are 0 and 8?

   e. (Show child page A. Cover all but top row of boxes.) There are some balls inside this box (point) and some balls inside this box (point). How many balls are there altogether, in both boxes?

   f. (Repeat "e" but with row 2 uncovered).

   g. How many balls are there in these two boxes (point)? (Row 3 is uncovered.

   h. ("g" but with row 4 uncovered).

   i. (Show child sheet "B") Here are some written questions. Tell me the answers

   (3t2)

   (2t3)

   (8t0)

   (0t8)

7. a. How many are 3 plus 4 plus 1?

   b. How many are 5 and 5 and 5?

   c. Place groups of 3 checkers, 4 checkers, ..and 1 checker on the table.) How many checkers are there in all?

   d. (Repeat "c" with 3 groups of 5 checkers each.)

8. I am going to count out loud. When I stop counting, tell me what number should come next.
   a. 1, 2, 3, --

   b. 1, 2, 3, 4, 5, --

9. Here are some (blue) beads (put them on table) and here are some (red) beads (put them on table). How many beads are there in all?

Child's name

10. 8 minus 5 is how many?

9 minus 4 is how many?

Here are some checkers (put 8 checkers on the table). How many are there? Now, take away 5 checkers. How many checkers are left?

(Repeat with a group of 9 checkers and have child remove 4)